மாதிரி அளவெடுப்புகளும் முறைகளும்

(SAMPLING SURVEYS AND METHODS)

கே. ஆர். சந்தானகோபாலன்

தமிழ்நாட்டுப் பாடநூல் நிறுவனம்

த.பா.நி. (க.வெ.) வரிசை எண் — 837

மாதிரி அளவெடுப்புகளும் முறைகளும்

ஆசிரியர்
கே. ஆர். சந்தானகோபாலன், எம். எஸ்ஸி.
புள்ளியியல் துணைப் பேராசிரியர்,
பூ.சா.கோ. கலைக் கல்லூரி,
கோவை.

தமிழ்நாட்டுப் பாடநூல் நிறுவனம்

First Edition—February 1979

Number of Copies—2000

T. N. T. B. S. (C. P.) No. 837

SAMPLING SURVEYS AND METHODS

K. R. SANTHANAGOPALAN

Price Rs. 12.30

Published by the Tamilnadu Textbook Society under the Centrally Sponsored Scheme of production of books and literature in regional languages at the University level, of the Government of India in the Ministry of Education and Social Welfare (Department of Culture), New Delhi.

This book has been printed on concessional paper made available by the Government of India.

Printed by

MUTHUKUMARAN PRESS,
13, Kuppier Street,
Madras—600,001.

அணிந்துரை

(திரு. செ. அரங்கநாயகம், தமிழகக் கல்வி அமைச்சர்)

தமிழைக் கல்லூரிக் கல்விமொழியாக ஆக்கிப் பதினெட்டாண்டுகள் ஆகிவிட்டன. குறிப்பிட்ட சில கல்லூரிகளில் இளங்கலை வகுப்புவரை மாணவர்கள் தங்கள் பாடங்கள் அனைத்தையும் தமிழிலேயே கற்று வந்தனர். 1969ஆம் ஆண்டிலிருந்து அறிவியல் பாடங்களையும் தமிழிலேயே கற்பிக்க ஏற்பாடு செய்துள்ளோம். தமிழிலேயே கற்பிப்போம் என முன் வந்துள்ள கல்லூரி ஆசிரியர்களின் ஊக்கம், பிற பல துறைகளில் தொண்டு செய்வோர் இதற்கெனத் தந்த உழைப்பு, தங்கள் சிறப்புத் துறைகளில் நூல்கள் எழுதித்தர முன்வந்துள்ள நூலாசிரியர்கள் தொண்டுணர்ச்சி இவற்றின் காரணமாக இத் திட்டம் நம்மிடையே மகிழ்ச்சியும் மனநிறைவும் தரத்தக்க வகையில் நடைபெற்று வருகிறது. இவ்வகையில் கல்லூரிப் பேராசிரியர்கள் கலை, அறிவியல் பாடங்களை மாணவர்களுக்குத் தமிழிலேயே பயிற்றுவிப்பதற்குத் தேவையான பயிற்சியைப் பெறுவதற்கு மதுரைப் பல்கலைக்கழகமும் சென்னைப் பல்கலைக் கழகமும் ஆண்டுதோறும் எடுத்துவரும் பெருமுயற்சியைக் குறிப்பிட்டுச் சொல்லவேண்டும்.

வரலாற்றியல், அரசியல், உளவியல், பொருளியல், மெய்ப் பொருளியல், புவியியல், புவியமைப்பியல், மனையியல், கணித வியல், இயற்பியல், வேதியியல், உயிரியல், வானியல், புள்ளி யியல், விலங்கியல், தாவரவியல், பொறியியல், சட்டவியல் ஆகிய எல்லாத் துறைகளிலும் மூலநூல்கள், மொழிபெயர்ப்பு நூல்கள் என்ற இருவகையிலும் தமிழ்நாட்டுப் பாடநூல் நிறுவனம் நூல்களை வெளியிட்டு வருகிறது.

இவற்றுள் ஒன்றான மாதிரி அளவெடுப்புகளும் முறைகளும் என்னும் இந் நூல் தமிழ்நாட்டுப் பாடநூல் நிறுவனத்தின் 837 ஆவது வெளியீடாகும். கல்லூரித் தமிழ்க் குழுவின் சார்பில் வெளியான 35நூல்களையும் சேர்த்து இதுவரை 872நூல்கள் வெளி வந்துள்ளன. இந்நூல் மைய அரசு, கல்வி, சமூகநல அமைச் சகத்தின் 'மாநில மொழியில் பல்கலைக்கழக நூல்கள் வெளியிடும் திட்டத்'தின்கீழ் வெளியிடப்படுகிறது

தமிழில் பயிலும் மாணவர்கள் உலக மாணவர்களிடையே சிறந்த இடம் பெறவேண்டும் என்பதே நம் குறிக்கோளாகும். கல்லூரிகளிலும் பல்கலைக்கழகங்களிலும் கலையியற் பாடங்களையும், அறிவியற் பாடங்களையும், தொழில்நுட்ப அறிவுப் பாடங்களையும் பயிலுகின்ற மாணவர்கள், அவற்றைத் தமிழில் பயில வேண்டும் என்பதை வலியுறுத்தி வருவதற்குக் காரணம், தமிழறிவு வளர வேண்டும் என்பதைவிட, தமிழ் மக்களின் அறிவு ஆற்றல் எளிதாக, விரைவாக வளரவேண்டும் என்பதுதான். 'எதிலும் தமிழ்; எங்கும் தமிழ்' என்னும் குறிக்கோளை நிறைவேற்ற வேண்டிய கடப்பாடு தமிழக ஆசிரியப் பெருமக்களையும் மாணவர்களையும் சார்ந்ததாகும். தமிழ்நாட்டுப் பல்கலைக்கழகங் களின் பல்வகை உதவிகளுக்கும் ஒத்துழைப்புக்கும் நம் மனம் கலந்த நன்றி உரித்தாகுக!

செ. அரங்கநாயகம்

நூன்முகம்

அன்னைத் தமிழ்மொழியில் பல்வேறு நூல்கள் வெளியிட்டு வரும் 'தமிழ்நாட்டுப் பாடநூல் நிறுவன'த்தின் திட்டத்தில், எளியேனது சிறு காணிக்கையாக, புள்ளியியல் கற்கும் மாணவர்கட்கு உதவும்பொருட்டு, 'மாதிரி அளவெடுப்புகளும் முறைகளும்' (Sampling Surveys and Methods) என்னும் இந் நூல் அமைகிறது

தற்காலத்தில், நாட்டின் பல்வேறு துறைகளிலும் உள்ள பிரச்சினைகளை ஆராயவும், அவற்றைத் தீர்ப்பதற்கான முடிவுகளைப் பெறவும், புள்ளியியல் செயல் முறையான 'மாதிரி அளவெடுப்பு' என்னும் வழிமுறை பெரும் அளவில் உதவுகின்றது. முழுமைத் தொகுதியின் பல்வேறு இயல்புகளை அறிய, குறைந்த முயற்சி, செலவு மற்றும் கால அளவு ஆகியவற்றுடன் திகழும் மாதிரி ஆய்வு வழிமுறைகள் பற்றிய விரிவான கருத்துகள் ஆராயப்படுகின்றன. முதல் மூன்று அத்தியாயங்களில், மாதிரி ஆய்வு முறை (Sample Survey Method) பற்றிய அடிப்படையான, பொதுவான விவரங்கள் கொடுக்கப்பட்டுள்ளன.

அதன் பின்னர், மாதிரித் தேர்வில் கருதப்படும் 'ராண்டம் மாதிரி முறை', 'படுகை மாதிரி முறை', 'ஒழுங்கு மாதிரி முறை', 'திரள் மாதிரி முறை', 'PPS மாதிரி முறை' போன்ற மிக முக்கியமான வழிமுறைகள் விவரமாக ஆராயப்பட்டுள்ளன. மேலும், ஒவ்வொரு மாதிரி முறையிலும் கணித அடிப்படையில் அமையும் முக்கியமான முடிவுகளும் தரப்படுகின்றன. விகித மதிப்பீடு மற்றும் தொடர்பு மதிப்பீடு ஆகியவை பற்றியும் விவரிக்கப்படுகின்றன.

ஒவ்வொரு பகுதியிலும் அறிமுறை நியதிகள் எளிதில் புரிந்து கொள்ளப்படும் விதத்தில், நிறைய எடுத்துக்காட்டுகளும் செய்முறைப் பயிற்சிகளும் இணைக்கப்பட்டுள்ளன.

மாதிரித் தேர்வு முறை சாராத பிழைகள், பதிலின்மை விளைவுகள் ஆகிய கருத்துகள்பற்றியும் விவரமாகத் தரப்பட்டுள்ளன.

நடைமுறைகளில் மேற்கொள்ளப்படும் மாதிரிக் கணிப்பாய்வுகளுக்கு இக் கருத்துகள் உறுதுணையாவன.

அடுத்து, நம் நாட்டில் பல்வேறு துறைகளிலும் மேற்கொள்ளப்படும் புள்ளிவிவர ஆய்வுகளையும், புள்ளிவிவரங்களின் தன்மைகளையும், 'இந்திய நாட்டுப் புள்ளிவிவரங்கள்' (Indian Statistics) என்ற பொதுவான தலைப்பின்கீழ் விரிவாகத் தரப்படுவதைக் காணலாம். இவை, இந்திய நாட்டில் நிலவும் சூழ்நிலைக் கேற்பப் பல்வேறு புள்ளிவிவரங்கள் அமையும் தன்மையை மாணவர் எளிதில் புரிந்துகொள்ளப் பெரிதும் உதவும்.

இறுதியாக, தற்போது பல்கலைக்கழக அளவில் அதிமுக்கியத்துவம் அளிக்கப்பட்டுவரும் 'வினா வங்கி' (Question Bank) என்ற பகுதியும் இணைக்கப்பட்டு, பல்வேறு தலைப்புகளிலும் வினாக்கள் சேகரிக்கப்பட்டு, தொகுத்து அளிக்கப்பட்டுள்ளன. பெரும்பாலும், பல்கலைக்கழகத் தேர்வுகளில் இடம்பெற்ற வினாக்கள் திரட்டப்பட்டுள்ளன.

இந் நூலினைத் தமிழில் மூலநூலாக அமைத்திட எனக்குப் பலவகையிலும் உதவிபுரிந்த, எம்துறைப் பேராசிரியர்கட்கு என் மனமார்ந்த நன்றியைத் தெரிவித்துக் கொள்கிறேன்.

இப் பணியை எனக்கு அளித்த, 'தமிழ்நாட்டுப் பாடநூல் நிறுவன'த்தினருக்கு என் மனப்பூர்வமான நன்றியைத் தெரிவித்துக்கொள்கிறேன்.

வாழ்க தமிழ் !

கே. ஆர். சந்தானகோபாலன்

பொருளடக்கம்

1. மாதிரி அளவெடுப்பு-ஓர் அறிமுகம்
(Sample Survey - An Introduction)

[புள்ளி விவரங்களின் தேவை—விவரங்களின் தன்மைகள்—முழுக் கணிப்பாய்வு—மாதிரி அளவெடுப்பின் தேவை]

1. புள்ளி விவரங்களின் தேவை (Need for Statistical Data)

மாதிரிக் கணிப்பாய்வு பற்றிய பல்வேறு இயல்புகளைக் கண்டறியுமுன், தற்காலச் சூழ்நிலையில், எந்த ஒரு நாட்டிலும் சேகரிக்கப்படும் பலவகைப்பட்ட புள்ளி விவரங்களின் இன்றியமையாமையினைப்பற்றிய ஒருசில கருத்துகளை முதன்முதலில் கண்டறிவோம். அறிவியல் மற்றும் தொழில்நுட்பத் துறைகளில் பல்வேறு வியக்கத்தக்க முன்னேற்றங்கள் ஏற்பட்டுவரும் இந் நூற்றாண்டினில், மக்களின் பொருளாதார மற்றும் சமூக வாழ்வு, போக்குவரத்து மற்றும் தகவல் தொடர்பு, தொழில் மற்றும் வணிகம், கல்வி மற்றும் மருத்துவ வசதிகள் போன்ற பலதரப்பட்ட இயல்புகளையும் தெளிவாக அறியவேண்டுவது மிக இன்றியமையாததாகிறது. நாடுகளில் பெரும் அளவில் உற்பத்தி மேற்கொள்ளப்பட்டு, குறைந்த செலவில் நிறைந்த பயன்களைப் பெறும் அளவிற்குத் திட்டங்கள் தீட்டப்படுகின்றன.

அரசுத் துறையினைச் சார்ந்து, அல்லது தனியார் துறையினில் அமைந்து விளங்கும் தொழில் நிறுவனங்கள், கூட்டமைப்புகள், வர்த்தகக் குழுமங்கள் போன்ற பலவும் பெரிய அளவினில் முயற்சிகளை மேற்கொண்டு, சரியானபடி திட்டமிட்டு, திறமையுடன் முடிவுகளைப் பெற்றுப் பயனடைய, கிடைக்கும் மூலப் பொருள்களையும், நிலவும் சூழ்நிலைகளையும் பற்றிய தெளிவான விவரக் கூறுகள் மிகவும் தேவைப்படுகின்றன. எனவே, குறிப்பிட்ட நோக்கத்திற்கு ஏற்ற வகையில் வெவ்வேறு வகையான புள்ளி விவரங்களைத் திரட்டி, தகுந்த முறைகளின் மூலம் பகுப்பாய்வு நடத்தி, இறுதி முடிவுகளின் அடிப்படையில் எதிர்காலத் திட்ட முறைகள் வகுக்கவேண்டியதாகிறது. தற்காலத்தில், அரசினர் தொழிற்கூடங்கள், வர்த்தக அமைப்புகள், பொது நிர்வாகங்கள் போன்ற பலருக்கும் புள்ளி விவரங்கள் அதிகமாகத் தேவைப்படுகின்றன.

நாட்டின் ஒழுங்கு மற்றும் கட்டுப்பாடு இவற்றைக் கவனிப்ப தோடு அரசின் பணி நின்றுவிடுவதில்லை. அதன் மக்களின் பொருளாதார, சமூக முன்னேற்றங்களுக்குத் தேவையான வழி முறைகளை வகுப்பதும் அதன் கடமையாகிறது. உதாரணமாக, நாட்டின் மக்கள்தொகைப் பெருக்கக் கட்டுப்பாடு, அதிக அளவில் உற்பத்தியைப் பெருக்குதல், அதிக அளவில் அடிப்படை வசதி களை அனைவருக்கும் அளித்தல், விலைவாசிகளைக் கட்டுப்படுத்தல் போன்ற பல்வேறு பொறுப்புகளைக் கூறலாம். தக்க அளவில் பொருளாதார முன்னேற்றம் அடையத்தக்க விதத்தில், மிக்க கவ னத்துடன் திட்டங்கள் வகுக்கப்படுகின்றன. மக்கள் நலனைப் பெரிதும் கருதி, வாழ்க்கைத் தரத்தினை உயர்த்தப் பெரும் முயற்சி எடுக்கப்படவேண்டும். இத்தகைய பல்வேறு பொறுப்புகளுக்கும் மிகவும் தேவைப்படுவன பல்வேறு துறைகளைச் சார்ந்த புள்ளி விவரங்களே ஆகும்.

நாட்டுப் பொருளாதாரத்தின் பல்வேறு உட்கூறுகள் (sectors) பற்றியும் விரிவான அளவில் விவரங்கள் சேகரிக்கப்பட வேண்டி யது அவசியம். மேலும், நாட்டின் வெவ்வேறு வகை நிலப் பரப்புகள், வெவ்வேறு விதமான மக்கள், வெவ்வேறு சூழ்நிலைகள், பழக்கவழக்கங்கள், வாழும் முறைகள் போன்றவற்றோடு, பயிர் செய்ய அமைந்திருக்கும் நிலப்பரப்பு, கிடைக்கும் மூலப் பொருள் கள், தாதுப்பொருள்கள், எண்ணெய் வளங்கள் போன்ற பல விவரங்களும் அதிகம் தேவைப்படுகின்றன. கிடைக்கும் மூலப் பொருள்களும் வசதிகளும் ஒரு வரம்புக்கு உட்பட்டு அமையும் நிலையில், அவற்றைப் பயன்படுத்துவது அதிகத் திறமையுடனும், கவனத்துடனும், எல்லாரும் எல்லாமும் பெறத்தக்க விதத்தில், இல்லாமை இல்லாத நிலை அமையும் வண்ணம் இருத்தல் இன்றி யமையாதது. ஆனால், மனிதத் தேவைகளோ அதிக அளவில் பெருகுகின்றன.

இந்நிலையில், ஓரளவு சமநிலை பெறத்தக்க விதத்தில், திறமை யாகத் திட்டங்கள் வகுக்கப்பட்டு, எதிர்காலத் தேவைகளையும் மனத்திற் கொண்டு செயல்பட, விரிவான அளவில், இருக்கும் சூழ் நிலைகளையும் வசதிகளையும்பற்றிய தெளிவான விவரங்கள் உதவு கின்றன.

பல்வேறு பொருள்களின் உற்பத்தி மற்றும் நாட்டின் தேவைக் குப் பொருள்கள் பயன்படும் அளவு ஆகியவற்றைப்பற்றி விவ ரங்கள் பெறுவது, ஏற்றுமதி-இறக்குமதிபற்றிய திட்ட வடிவமைப் பிற்கு மிக உதவும். ஊதிய விகிதம், விலைவாசிகள் ஆகியவற்றை

நிர்ணயிக்க, வெவ்வேறு பிரிவு மக்களின் வாழ்க்கைச் செலவுக் குறியீட்டெண் பற்றிய விவரங்கள் அடிப்படையாக அமை கின்றன. பிறப்பு இறப்பு விவரங்கள், நிலவும் நோய்கள்பற்றிய விவரம், கல்வி அளவு பற்றிய தெளிவு, வாழும் சூழ்நிலை பற்றிய அறிவு போன்ற பலவகையான புள்ளி விவரங்களும் வெவ்வேறு தன்மையுடைய ஆய்வுகளில் பெரிதும் பயன்படுகின்றன. இவை நாட்டின் பொருளாதார மற்றும் சமூக முன்னேற்றத்திற்குப் பெரிதும் வேண்டப்படுவன ஆகும்.

சுருக்கமாக, புள்ளி விவரங்கள் மனித வாழ்வின் இன்றியமை யாத அடிப்படையாக அமைகின்றன என்பது மிகையாகாது. மேலும், தொழில், வர்த்தகத் துறைகளில், செலவு, மொத்த இருப்பு, உற்பத்தி அளவு, நுகர்வோர் தேவைகள், இலாப விகிதம் போன்ற பல விவரங்களின் தெளிவு, விற்பனையினைப் பெருக்கவும் உற்பத்தியை அதிகரிக்கவும் தேவையான திட்டங்களை வகுக்கப் பெரும் அளவில் உதவிபுரிகின்றன. எனவே, எந்தத் துறையிலும் புள்ளி விவரங்கள் சேகரிப்பதானது ஒரு மிக முக்கிய இடத்தைப் பெற்றுத் திகழ்கிறதெனலாம்.

2. விவரங்களின் தன்மைகள் (Nature of Data)

பல்வேறு நோக்கங்களுக்குப் பயன்படுத்தப்படும் புள்ளி விவரங்கள் இருவகையாக அமைகின்றன :

(i) ஆய்வு விவரங்கள் (Survey data).

(ii) பரிசோதனை விவரங்கள் (Experimental data).

(i) நாம் பிற்பகுதிகளில் விரிவாகக் கருதும் ஆய்வு முறைகள் முதல் வகையினைப் பற்றியே அமைகின்றன. அத்தகைய விவரங்கள் முறையாகக் கணிப்பாய்வுகள்மூலமாகப் பெறப்பட்டுப் பயன்படுத்தப்படுகின்றன. இத்தகைய ஆய்வு விவரங்களைப் பின் வருமாறு கருதலாம் :

வெவ்வேறு இயல்புகளும் (characteristics) குறிப்பிட்ட ஒரு காலத்தினை அல்லது கால இடைவெளியினைக் கருதிப் பெறப்பட லாம். உதாரணமாக, மக்கள்தொகைபற்றிய விவரம், பயிர்ச் சாகு படி அளவு, தொழில் உற்பத்தி விகிதம், பல்வேறு தொழில்களில் ஈடுபட்டிருப்போர் எண்ணிக்கை, மக்கள்தொகை நெருக்கம். போன்ற பலதரப்பட்ட விவரங்கள் இதன்படி பெறப்படுகின்றன. இதோடு, நுகர்வோரின் விழைவு, வேலையின்மை விவரம், கல்வித் தரம், மக்கள் தொகை அதிகரிக்கும் வீதம் போன்ற பலவகை

விவரங்களும் சுட்டிக் காட்டப்படலாம். இவை அனைத்தும் காலத்தோடு தொடர்பு கொண்டு மாறும் தன்மையுடன் அமைவதால், காலத் தொடர்புடன்கூடிய விவரங்களை அடிக்கடி பெறவேண்டியதாகிறது.

அடுத்த வகையான ஆய்வு விவரங்கள், காலத்தோடு ஒட்டி வெவ்வேறு இயல்புகளும் மாறும் போக்கினையும் (trend), ஏற்படும் தொடர்புகளையும் பற்றியதாக அமைகின்றன. கண்டறியப்பட்ட விவரங்களின் அடிப்படையில், எதிர்காலத் தேவைகளும் வழிமுறைகளும் நிர்ணயிக்கப்படுகின்றன. இதோடுகூட, அரசினர் மேற்கொள்ளும் நடவடிக்கைகள் எந்த அளவில் மக்களது வாழ்வினைப் பாதிக்கின்றன என்றும் அறிய முடிகிறது. எடுத்துக்காட்டாக, மக்களின் வாழும் சூழ்நிலைக்கும், பொருளாதார மற்றும் சமூகக் கூறுகட்குமிடையில் தொடர்பினைக் கண்டறிவதன்மூலம், எதிர்காலத்தில் தேவைப்படும் உணவுப் பொருள்கள், எரிபொருள் துணி, வீட்டு வசதி, மருத்துவ வசதி போன்ற பலவற்றையும் மதிப்பிட முடிகிறது.

அடுத்த வகை விவரங்கள் நாட்டின் நில அமைப்பு, மற்றும் பூகோள அமைப்பு பற்றி அமைகின்றன. மண்ணின் தன்மை (soil type), தட்பவெப்பநிலை (climatic conditions), கிடைக்கும் தாதுவளம், எண்ணெய் வளம் போன்ற பலவும் இவற்றில் அடங்கும். இவற்றைப்பற்றி அறியப்படும் விவரங்கள் எல்லா வகையான ஆய்விற்கும் மிகப் பயன்படும் என்பது குறிப்பிடத் தக்கதாகும்.

(ii) இனி, இரண்டாம் வகையான பரிசோதனை விவரங்கள், தகுந்தத் திட்ட வரையினை மேற்கொண்டு நடத்தப்படும் புள்ளியியற் சோதனைகள்மூலம் திரட்டப்படுவன ஆகும். அவை இந்நூலின் நோக்கத்திற்குப் புறம்பாக அமைவதால், பயிர்கள் பற்றிய ஆய்வுகளில், குறிப்பாக விவசாயத் துறையில் அவை அதிகம் பயன்படுத்தப்படுகின்றன என்றுமட்டும் கூறி அவற்றை இந்த அளவில் விட்டுவிடுவோம்.

மக்கள்தொகை, பிறப்பு-இறப்பு விவரங்கள், வரவு-செலவு விவரங்கள், விலைவாசிகள், வீட்டுவசதி நிலைகள், முதலீட்டு உரு அமைப்பு, ஆக்கப்பணிகள், கல்வி, மருத்துவ நிலைகள், நில உடைமையும் பயனும் பற்றிய உண்மைகள், உற்பத்திச் செலவு, போக்குவரத்து, தகவல் தொடர்பு, அரசின் நடவடிக்கைபற்றி மக்களது கருத்துகள், பொருள்களின் தேவைபற்றிய சந்தை

நிலைகள் (markets), விற்பனை அளவுகள் போன்ற பல்வேறு வகைகளில் அமைந்த விவரங்களைத் திரட்ட, விரிவான அளவில் ஆய்வுகள் மேற்கொள்ளப்படுகின்றன. இந்த ஆய்வுகள் திறனுடன் மேற்கொள்ளப்படின், முடிவுகளும் சிறந்து விளங்கும்.

நாட்டின் பொருளாதாரத் திட்டங்களுக்கு உதவும் வகையில் விவரங்கள் சேகரிக்கப்படும். உதாரணமாக, மக்கள்தொகைக் கணிப்புகளில் பெறப்படும் பல்வேறு உண்மைகள் அரசின் பல் வேறு திட்டப் பணிகளுக்குப் பேருதவி புரிகின்றன. இதேபோல் தொழில், வர்த்தகம் போன்ற துறைகளும் ஆய்வு முடிவுகளை அதிகம் பயன்படுத்துகின்றன.

3. முழுக்கணிப்பு முறை (Census Method)

ஒரு முழுமைத் தொகுதியில் (population) அமைந்த எல்லா உறுப்பினர்களையும் அல்லது கூறுகளையும் (units) கருதி, விவரங்களைப் பெறுவது 'முழு ஆய்வு முறை' (complete enumeration) எனப்படும். இதற்கான காலம், செலவு, முயற்சி ஆகியவை பெரும் அளவினதாக இருக்கும். எனவே, மிகவும் தேவைப்படும் இடங்களில் மட்டுமே இம்முறையினைப் பயன்படுத்த வேண்டும். மிகவும் சரியாக, துல்லியமான (accurate) விவரங்கள் தேவைப்படும் பொழுது, இம் முறை ஆய்வானது தவிர்க்கப்பட இயலாது. உதாரணமாக, மக்கள்தொகைக் கணக்கெடுப்பு (population census), வருமான வரிக் கணக்கீடு (income tax assessment) போன்றவற்றில் இம் முறையே பொருத்தமாக அமைகிறது. ஆய்வின் அடிப்படை நோக்கத்தையும், இறுதி முடிவின் தன்மையையும் (objective and purpose) பொறுத்து இம்முறை பின்பற்றப்படல் வேண்டும். இதிலும், முற்றிலும் பிழையற்ற கணிப்பு அமைவது இயலாது.

பொதுவாக, சில மதிப்பீடுகளில் ஓரளவு அனுமதிக்கப்படக் கூடிய பிழையுடன் (permissible error) விவரங்கள் குறிக்கப்படுகின்றன. உதாரணமாக, வெவ்வேறு இடங்களின் மக்கள்தொகைகளை ஒப்பிடுகையில், முழு எண்களைக் கொண்டே ஒப்பீடு (comparison) நடைபெறுகிறது. இதேபோல், இந்த ஆண்டு நாட்டு மக்களின் தேவைக்கு 200 மிலியன் டன்கள் உணவுப் பொருள் தேவைப்படும் (உதாரணமாக) என்று கூறுவது, சென்ற ஆண்டிற்குப் பெறப்பட்ட மதிப்பீடான 196 முதல் 204-க்குள் அமையும் அளவினைப் பொறுத்து அமைந்திருக்கும். எனவே, இதன் அடிப்படையில் ஏற்றுமதி பற்றிய விதிமுறையை (export policy) திட்டமிடலாம். அதாவது $\pm$ 4 என்ற பிழை அளவில்

மதிப்பிடப்பெறும் உற்பத்தி அளவு அடிப்படையாகக் கொள்ளப் படும். இதைப் பொறுத்து முழுக் கணிப்பு தேவையா இல்லையா என முடிவு செய்யலாம்.

முழு அளவில் கணிப்பு நடைபெறும் ஆய்வுகளில் பிழை ஏற் படாது என்று உறுதி அளித்தல் இயலாது. உதாரணம், மக்கள் தொகையானது முற்றிலும் சரிவரக் கணக்கெடுக்கப்படுவது இய லாத காரியம். எனவே, முழு ஆய்வு முறையில், கருதப்பட்ட இடங்களில் ஆய்வு முழுமை பெருததாலும், கண்டறிவதில் ஏற் படும் தவறுகளாலும், அட்டவணைப்படுத்துதல் சரியாக அமை யாததாலும் பெரும் அளவில் பிழைகள் ஏற்பட நிறைய வாய்ப்பு உண்டு. எனவே, மிகவும் இன்றியமையாத சூழ்நிலைகளிலன்றி, இம்முறையானது பின்பற்றத் தக்கதன்று எனலாம்.

4. மாதிரி அளவெடுப்பின் தேவை

மிகவும் பெரிய அளவினதான ஒரு முழுமைத் தொகுதியின் இயல்புகளை ஆராய்ந்து கண்டறிவதற்கும், அதன் மூலம் சில பொதுப்படையான முடிவுகளை அடைதற்கும், மாதிரி அள வெடுப்பு அல்லது ஆய்வு முறைகள் மிகவும் உதவுகின்றன. துல்லியமான மதிப்பீடுகளைக் கணிப்பதற்கும், ஆராய்ந்து கண் டறியும் கால அளவினை மிகவும் குறைத்திடவும், இத்தகைய கணிப்புக்கு ஆகும் செலவைக் கட்டுப்படுத்தவும் மாதிரி முறை விசாரணை மிக இன்றியமையாததாகிறது. தொழிற் துறைகளி லும் மற்றும் பல்வேறு பொது முறைகளிலும் சிறந்த முடிவுகளைப் பெறுவதற்கு மாதிரி அளவெடுப்பு முறைகள் வழி செய்கின்றன. இம் முறைகள் பின்பற்றுதற்கும் எளிதாக அமைகின்றன. இதில் குறிப்பிடத்தக்க தென்னவெனில், கண்டறிந்த விவரங்களின் தன்மைக்கு ஏற்ற வகையில், ஆய்வு முறைகள் அமைந்திட வேண்டும்.

மாதிரி என்பது ஒரு முழுமைத் தொகுதியில் இருந்து எடுக் கப்பட்ட கூறு ஆகும். எனவே, முழுமைத் தொகுதியின் இயல்பு களைப் பிரதிபலிக்கத்தக்க வகையில் மாதிரியானது அமைதல் அவசியம். மக்கள் அல்லது பொருள்கள் நிரம்பிய ஒரு தொகுப் பினை முழுமைத் தொகுதி என வரையறுக்கலாம். உதாரணமாக, ஒரு நகரில் வாழும் மக்கள், ஒரு தொழிலகத்தில் உற்பத்தி செய் யப்படும் பொருள்கள், ஒரு நிலத்தில் விளைந்திட்ட தானிய அளவு போன்றவற்றை முழுமைத் தொகுதிகளாகக் கொள்ளலாம். கருதப்பட்ட ஒரு முழுமைத் தொகுதியின் இயல்புகள்பற்றிய மதிப்பீடுகள், முடிந்த அளவுக்கு உண்மை நிலையைக் காட்டும்

வண்ணம் தேவைப்படுகின்றன. ஒரு குடும்பத்தின் சராசரி மாத வருமானம், பிறப்பு விகிதத்தில் ஆண்களின் எண்ணிக்கை, சராசரியாக 1000 மணி நேரம் எரிந்திடும் மின் விளக்குகளின் எண்ணிக்கை போன்றவற்றைக் கணக்கிடுவதாகக் கொண்டால், அதற்கேற்பச் சில வழிமுறைகளை வகுக்க வேண்டியதாகிறது.

தகுந்த வழிமுறைகளுடன் மாதிரி ஆய்வானது மேற்கொள்ளப்படின், கருதப்பட்ட இயல்பின் (characteristic) மதிப்பீட்டினை (முழுமைத் தொகுதிபற்றி) சிறந்த முறையில் பெறவும், அம் மதிப்பீட்டின் மாதிரி சார்ந்த மாறுபாட்டினை (sampling variance) அறியவும் வழி ஏற்படுகிறது. இவற்றை, வெவ்வேறு மாதிரி முறைகளை விரிவாகக் கருதுகையில், விவரமாகப் பகுத்து ஆய்வோம். நடைமுறையில் மேற்கொள்ளப்படும் ஒரு வழிமுறையே மாதிரி முறையாக (sampling) உருப்பெறுகிறது. சரக்கின் தன்மையைச் சோதிக்க, வணிகர்கள் பெரும் மூட்டைகளினின்று எடுக்கப்படும் ஒரு சிறு பகுதியை மாதிரியாகக் கொண்டு முடிவெடுக்கின்றனர். மேலும், அன்றாட வாழ்வில் மற்றொரு நிகழ்ச்சியான, 'ஒரு பானை சோற்றுக்கு ஒரு சோறு பதம்' என்னும் பழமொழிக்கு ஏற்ப அமையும் ஒரு நிலையினையும் குறிப்பிடலாம். மாதிரி முறை பயன்படுத்தப்படாத இடமே இல்லை என்பதைக் கண்கூடாக அறியலாம்.

டெமிங் (W. E. Deming) கூறுவதுபோல், மாதிரி முறை ஆய்வானது, 'பயன்படுத்தக்கூடிய புள்ளிவிவரங்களின் நம்பகத் தன்மையினை அளவிட்டுக் கட்டுப்படுத்த, வாய்ப்பு விதிகளைக் கொண்டு அமைந்த' அறிவியலாகும்.

பல்வேறு துறைகளிலும் சிறந்த முடிவுகளைப் பெற மாதிரி ஆய்வு முறை பயன்படுத்தப்பட வேண்டியதன் தேவையினைப் பின்வரும் ஒப்பீடு மூலமாகத் தெள்ளத்தெளிய உணரலாம். அதாவது, முழுக்கணிப்பு முறையுடன் ஒப்பிடுகையில், மாதிரிக் கணிப்பாய்வு முறை எந்த விதத்தில் சிறந்து விளங்குகிறது என்பதனை விளக்கமாகக் காணலாம்.

முழுக் கணிப்பு, மாதிரி அளவெடுப்பு—ஓர் ஒப்பீடு (Complete Enumeration, Sampling—A Comparison)

மாதிரி முறை அளவெடுப்பின் சாதக மற்றும் பாதகக் கூறுகள் (merits and demerits) விரிவான அளவில், யேட்ஸ் (F. Yates), மஹலநோபிஸ் (P. Mahalanobis) போன்ற பல

புள்ளியியல் வல்லுனர்களாலும் சிறந்த முறையில் விவாதிக்கப் பட்டிருக்கின்றன. சிலவற்றைத் தொகுத்துப் பின்வருமாறு கருதலாம் :

முதன் முதலில், ஏதேனும் பொருள்பற்றி ஆய்வு மேற் கொள்கையில், சில அடிப்படை வசதிகள் நிறைந்த சூழ்நிலை இருக்குமென எதிர்பார்க்கப்படுகிறது. அதாவது, தேவையான நிதி வசதி, ஆய்வு நடத்தத் தேவைப்படும் பணியாளர்கள், இடப் பரப்புகளைக் காட்டும் போதிய வரைபடங்கள் (maps), போக்கு வரத்து வசதிகள் போன்ற பலவும் அமையும் நிலையில் ஆய்வு நடத்துவது எளிதாகிறது. இவை அனைத்தும் முழுக் கணிப்பில் போதிய அளவு அமைந்து விளங்கப் பெறுவதில்லை. எனவே, இவை குறைந்த அளவே தேவைப்படும் வழிமுறையான மாதிரிக் கணிப்பாய்வு (sampling method) முறையானது நடைமுறைக்கு எளிதாகிறது.

சிறு நிலைகளில் (small units) கருதப்படும் ஆய்வுகளில், மாதிரிக் கணிப்பாய்வு முறை பயன்படாததாகிறது. உதாரண மாக, குறிப்பிட்ட கால இடைவெளிகளில் எடுக்கப்படும் மக்கள் தொகைக் கணக்கெடுப்புகளில் (census) வெவ்வேறு கிராமப் புறங்களில், விவசாயம், மக்கள்தொகை, கால்நடைகள் போன்ற விவரங்கள் பெறப்படுகின்றன. மாதிரி ஆய்வுகளில் இதே அளவு திட்பமாக (precise) முடிவுகளைக் கொடுப்பதில்லை என்ப தால், பொதுவாகச் சிறு நிலைகளில் கருதப்படும் மாதிரிக் கணிப்பு அவசியமற்றதாகிறது. அதாவது, பெரிய அளவில் பிரிவுகளை (கிராமங்களின் தொகுதி போன்று) கருதி, மாதிரி ஆய்வு பயன் படுத்தப்படுகையில், முடிவுகள் திறன்பெற்று விளங்கும்.

சில ஆய்வுகளில் முழுக் கணிப்பானது அர்த்தமற்றதாகவும் அமைய நேரிடும். உதாரணமாக, உற்பத்தியாகும் ஒரு மின் விளக்கின் சராசரி ஆயுளைக் கணக்கிடுவதற்கு, உற்பத்தி செய் யப்பட்ட விளக்குகள் அனைத்தையும் ஒருங்கே பயன்படுத்துவது பொருத்தமற்றதாக அமைந்து, பெரும் அளவில் இழப்பையும் (loss) ஏற்படுத்தும். எனவே, இத்தகைய சூழ்நிலைகளில், முழுமைத் தொகுதியின் ஓர் உட் பிரிவினை (sub-unit) தகுந்த முறையில் தேர்ந்தெடுத்து ஆய்வு நடத்துவதே அறிவுடைமை ஆகும்.

மேலும், முந்தைய அனுபவங்களின்படி, தக்க வழிமுறை களுடன் நன்கு திட்டமிடப்பட்டு அமையும் மாதிரி ஆய்வுகளில்

(sample surveys), பல்வேறு இயல்புகளைப்பற்றித் திட்பமான மதிப்பீடுகள் பெறப்படும் என்று கூறலாம். நடைமுறைக்கு ஏற்ற அளவில் முடிவுகள் பெறப்படுகின்றன.

அடுத்து, மிக முக்கியமானதாக அமைவது ஆய்வுச் சிக்கனமாகும் (economy). விவரங்களைப் பெறுவது, தொகுப்பது போன்ற பல வகையான செயல் முறைகளும் குறைந்த செலவுடன் மாதிரி ஆய்வு முறைகளில் அமைகின்றன. இதன்மூலம், பல்வேறு இயல்புகளையும்பற்றிய பரந்த ஆய்வு மேற்கொள்ளப் பட்டு, நோக்கமும் விரிவடைகிறது.

அதிக அளவில் ஆய்வாளர்களையும், கணக்கெடுப்போரையும் பயன்படுத்தி முழுக் கணிப்பாய்வு மேற்கொள்ளப்படுவதால், ஏற்படும் பிறழ்ச்சிகளையும், மாதிரி முறை சாராத பிழைகளையும் கட்டுப்படுத்துவது இயலாது போகிறது. பெறப்பட்ட முடிவுகளை மறுபரிசீலனை செய்வதும், மறு ஆய்வு மேற்கொள்வதும் பெரும் பாடாக அமையும். இதற்கு மாறாக, மாதிரி அளவெடுப்பில் இம் முறைகள் தக்க வழியில் குறைந்த செலவில் மேற்கொள்ள இயல்வதாக உள்ளது. எனவே, சிறு கூறுகளாகப் பிரிக்கப்பட்ட பகுதிகளில் முழு ஆய்வுமூலம் முக்கிய விவரங்களைக் கருதுவதும், மற்றவற்றிற்கு அவற்றின் முக்கியத்துவத்துக்கேற்ப பெரிய அளவில் பகுதிகளைக் கருதி முடிவுகள் பெறுவதும் சிறந்ததாகக் கருதப்படும்.

சுருங்கக் கூறின், மாதிரிக் கணிப்பு ஆய்வு முறையானது குறைந்த அளவு காலம், செலவு, முயற்சி ஆகியவற்றுடன், பெரிய இடப்பரப்பில் மேற்கொள்ளத் தக்கதாக விளங்கும் எனலாம். இதன்பொருட்டு, கருதப்படும் மாதிரியானது, முழுமைத் தொகுதியின் முழு இயல்பினையும் பிரதிபலிக்கத் தக்கதாக அமையுமாறு தேர்ந்தெடுக்கப்படுவது இன்றியமையாதது.

2. மாதிரி அளவெடுப்பில் இடம்பெறும் கருத்துகள் மற்றும் வரையறைகள்
(Concepts and Definitions in Sample Surveys)

[அலகு — முழுமைத் தொகுதி வகைகள் — மாதிரி அலகுகள் மாதிரிச் 'சட்ட' வரையறை — பிழைகள் — மாதிரிக் கணிப்பும் பிறழ்ச்சியும் — இயல்நிலைப் பரவல் — பிறழ்ச்சி விளைவுகள் — மாறுபாடு மற்றும் செலவுச் சார்புகள் — மதிப்பீட்டுத் திட்பம் — மாதிரித் தேர்வின் தத்துவ நியதிகள்]

மாதிரித் தேர்வு முறை மற்றும் மதிப்பீட்டு முறை ஆகியவற்றினை, மாதிரிக் கணிப்பாய்வுகளைப் பொறுத்த அளவில், விரிவாகக் கருது முன்னர், அடிக்கடிப் பயன்படும் சில கருத்துகளும், வரையறைகளும், சொற்களும் பற்றிய விளக்கமான ஆய்வு மேற் கொள்ளப்பட வேண்டியுள்ளது. அதன் பின்னர் மாதிரித் தேர்வு வழி முறைகள், பின்பற்றப்படும் செயல்முறைகள் ஆகியவற்றினை எளிதில் அறிந்துகொள்ள இயலும்.

1. அலகு அல்லது கூறு (Unit)

தனித் தன்மையுடையதாகவும், கூட்டு அமைப்பாகவும் இருக்கும் தன்மை பெற்று, விவரங்கள் சேகரிக்கப்பட்ட தேர்ந்தெடுக்கும் நபர் அல்லது பொருளானது 'கூறு' எனப்படுவது இயல்பு. தனியாக எளிதில் கண்டு கொள்ளத்தக்க விதத்தில் கூறுகள் அமையவேண்டும். உதாரணம்: குடும்ப நபர், வீடு, குடும்பம், பண்ணை, நிறுவனம், நாற்காலி போன்ற பல. ஆய்வு நோக்கத்திற்கேற்ப, கூறுகள் வரையறுக்கப்படும். பல கூறுகள் சேர்ந்து விளங்குவது ஒரு முழுமைத் தொகுதியாகக் கருதப்படும். முழுமைத் தொகுதிக் கூறுகளைக் கருதி, மாதிரிக் கூறுகள் (Sample Units) தேர்ந்தெடுக்கப்படும். முதலில், முழுமைத் தொகுதிபற்றிச் சில கருத்துகளைக் கருதுவோம். அதன் பின்னர், மாதிரிபற்றி விவரங்களை அறிதல் எளிது.

2. முழுமைத் தொகுதியின் பல்வேறு வகைகள்

முழுமைத் தொகுதியிலிருந்து ஒரு மாதிரியினைத் தேர்ந்து எடுக்குமுன் அந்த முழுமைத் தொகுதியின் குணாதிசயங்கள்

தெரிந்திருத்தல் அவசியம். பலவிதமான பண்புகள் நிறைந்த ஒரு முழுமைத் தொகுதியின் ஏதேனுமொரு மாறுபடும் இயல்பினை அறியும் பொருட்டு, அதற்குப் பொருந்தாத ஒரு முறையினைப் பின்பற்றி ஒரு மாதிரியினைத் தெரிவு செய்வோமே எனில், அம் முழுமைத் தொகுதியின் இயல்பினை அந்த மாதிரியானது முற்றிலும் பிரதிபலிப்பதாய் அமையாது. எனவே, ஒரு மாதிரியினைத் தேர்ந்தெடுக்கு முன்னர், அதன் முழுமைத் தொகுதி எந்த இனத்தைச் சேர்ந்தது என அறிய வேண்டும். முழுமைத் தொகுதி பற்றிய எந்த விவரமும் அறியாத நிலையில், மாதிரி அளவீட்டு முறை பயன் தராது.

முன்பு கண்டபடி, புள்ளியியல் முழுமைத் தொகுதி என்பது உயிருள்ள அல்லது உயிரற்ற பொருள்களின் தொகுப்பு என வரையறுக்கப்படுகிறது. அன்றாட வழக்கில் முழுமைத் தொகுதி (population), மக்கள் நிரம்பிய தொகுப்பைக் குறித்தாலும் புள்ளியியல் முறைப்படி பொருள்களின் முழுமை (universe) என்பதுதான் கருதப்படுகிறது. இம் முழுமைத் தொகுதியானது,

(i) வரம்புள்ள முழுமைத் தொகுதி (Finite Population),

(ii) வரம்பற்ற முழுமைத் தொகுதி (Infinite Population),

(iii) மெய்யான முழுமைத் தொகுதி (Real Population),

(iv) கற்பனை முழுமைத் தொகுதி (Hypothetical Population) என்று பகுக்கப்படுகிறது.

ஒரு மரத்தில் விளைந்த கனிகள், கல்லூரியில் பயிலும் மாணவர்களின் எண்ணிக்கை, ஒரு புத்தகத்தில் உள்ள பக்கங்கள் போன்றவை (i)-ல் அடங்கும். மாறாக, ஒரு கானகத்தில் உள்ள மரங்கள், பல்வேறு நேரங்களில் குறிக்கப் படும் தட்பவெப்ப நிலைகள் போன்றவை எண்ணிக்கையில் வரை யறுக்க முடியாததாக (ii)ஆவது வகையில் சேரும். உண்மையில் இயக்கம் உள்ள பொருள்கள் (iii)ஆம் வகையில் அடங்கும். உருவம், வடிவம் கொண்ட அனைத்துப் பொருள்களும் இதில் அடங்கும். இதற்கு முற்றிலும் மாறாக, கற்பனை செய்யப்படும் முழுமைத் தொகுதி அமைந்திருக்கும். உதாரணமாக, ஒரு நாணயத்தைக் கணக்கற்ற முறைகள் (5000-க்கு மேலாக) சுண்டிப்பார்க்கையில், அதன் விளைவுகள் (பூ அல்லது தலை) இத் தகைய ஒரு முழுமைத் தொகுதியைத் தரும்.

3. மாதிரி அலகுகள் அல்லது மாதிரிக் கூறுகள் (Sampling Units)

ஓர் அடிப்படை அலகு அல்லது அலகுகளின் தொகுதியானது மாதிரி எடுப்பில் மிகவும் தேவையான ஒன்று. இத்தகைய அலகுகள் அல்லது உறுப்புகள் பற்றிய புள்ளி விவரங்கள் ஒரு வரையறுத்த முறைப்படி பெறப்படுகின்றன. பல தருணங்களில், ஒரு முழுமைத் தொகுதியில் உள்ள சில பகுதிகள் மட்டும் தேவைப்படுகின்றன. உதாரணமாக, ஓர் ஆய்வில் நாம் கருதுவது, நன்கு தேர்ந்தெடுக்கப்பட்ட சில குடும்பங்களாக இருக்கலாம்.

ஒரு முழுமைத் தொகுதியில் அமைந்த அடிப்படை உறுப்புகள், மாதிரி உறுப்புகள் என்று கூறப்படுகின்றன. இவற்றை நன்கு வரையறுக்கவும், தனியாக அடையாளம் காணவும் இயலவேண்டும். எடுத்துக்கொண்ட பகுப்பாய்வுக்கு ஏற்ற விதத்தில் அலகுகள் தேர்ந்தெடுக்கப்பட வேண்டும். எடுத்துக்காட்டாக, குடும்ப வரவு-செலவுத் திட்டம் பற்றிய ஆய்வினில், ஒரு குடும்பத்தினை மாதிரி அலகாகக் கருத வேண்டும். பள்ளி மாணவரிடையே எடுக்கப்பட்ட மாதிரியில், ஒவ்வொரு மாணவனும் ஒரு **தனி** அடிப்படை அலகாக அமைகிறன். அல்லது மாதிரியானது ஓர் இடத்தில் அமைந்த பள்ளிகளில் இருந்து தெரிவு செய்யப்படின், ஒவ்வொரு பள்ளியும் ஒரு **தொகுப்பு (cluster)** அலகாக அமையும்.

மாதிரி அலகு ஒவ்வொன்றும் தனித் தன்மைத்தாகவும் மேற்படிதல் இல்லாமலும் (non-overlapping) அமைந்திருக்க வேண்டும். ஒவ்வோர் அடிப்படை அலகும் ஏதேனும் ஒரு மாதிரியில் தேர்ந்தெடுக்கப்படும் வாய்ப்புப் பெற்றதாக அமைய வேண்டும். ஒரு குறிப்பிட்ட வகையினதான மாதிரி அலகுகள் கொண்ட தொகுப்பினை மாதிரி அலகுகளின் முழுமைத் தொகுதி எனலாம். இத்தகைய அலகுகளை முன்கூட்டியே நன்கு வரையறுத்தல் மாதிரி அளவீடு பற்றிய பகுப்பாய்வில் மிகவும் இன்றியமையாத முதற்கட்டப் பணியாகும். இந்தக் கட்டத்துடன் தொடர்புடைய பின்வரும் சட்டத்தினை வரையறை செய்வோம்.

4. மாதிரி சட்ட வரையறை (Sampling Frame—Definition)

மாதிரி அலகுகளை (ஒரு முழுமைத் தொகுதியின் அடிப்படையில்), வரிசைப்படுத்திய ஒரு விரிவான பட்டியல் (list) மாதிரிச் சட்டம் என்று கூறப்படும். இத்தகைய சட்டத்தி

னின்று மாதிரி அலகுகள் தெரிவு செய்யப்படுகின்றன. அலகுகளை நன்கு தெரியப்படுத்தும் பட்டியல் அல்லது மாதிரி அலகுகளின் எல்லை வரையறையினைக் காட்டும் வரைபடம் இவ்வாறு அழைக்கப்படுகிறது. இத்தகைய பட்டியலின் அடிப்படையில் மாதிரியானது எடுக்கப்படுகிறது. கருதப்பட்ட முழுமைத் தொகுதி பற்றியே இது அமைய வேண்டும். நடைமுறைக்கு ஏற்றதாக, அதிக மாற்றங்களின்றி இருப்பின், இத்தகைய பட்டியல் மிகவும் பயன்தரும். வேறு சில விவரங்களும் தரப்பட்டிருப்பின், மாதிரித் தேர்வு முறை சிறந்ததாகவும், இறுதியில் பெறப்படும் முடிவுகள் நம்பகமானவையாகவும் இருக்க ஏதுவாகும். இத்தகைய பட்டியல்களுக்குப் பின்வரும் உதாரணங்களைச் சுட்டிக் காட்டலாம்:

(i) கோவை மாவட்டப் பஞ்சாலைத் தொழிலாளர் பற்றிய விரிவான பட்டியல்.

(ii) சென்னைப் பல்கலைக் கழகப் பரப்பில் அமைந்த கலைக் கல்லூரிகளில் பயிலும் மாணவர் பற்றிய முழுமைப் பட்டியல்.

(iii) 1971 மக்கள்தொகைக் கணிப்பின்படி அமைக்கப்பட்ட, இந்தியக் கிராமங்கள் பற்றிய நீண்ட பட்டியல்.

இவை தவிர மாதிரிக் கூறுகள் அமைந்த எல்லையை வரையறுக்கும் பூகோள அடிப்படை விளக்கப் படங்களும் மாதிரிச் சட்டத்தின் ஓர் அமைப்பு எனலாம். பொதுவாக, மாதிரிச் சட்டமானது, முழுமைத் தொகுதியின் ஒரு வடிவமே ஆகும். வரிசைப் படுத்தப்பட்ட விவரங்கள் மாதிரித் தேர்வினை எளிதாக்குகின்றன. இம் முறையில், முழுமையின்மை, விவரங்களின் பற்றுக்குறை நிலை, காலங் கடந்த விவரங்களைக் குறித்திடல் ஆகியவற்றை நீக்கத் தேவையான வகையில் நடவடிக்கை எடுக்கப்படல் வேண்டும். மாதிரி அலகுகள் மற்றும் மாதிரிச் சட்டம் இரண்டும் மாதிரி வழிமுறையின் அடிப்படைத் தேவைகள்.

5. பிழைகள் (அல்லது வழுக்கள்) (Errors)

இவை இரு வகையாகப் பிரிக்கப்படுகின்றன:

(1) மாதிரி முறை சார்ந்த பிழை (Sampling Error)

(2) மாதிரி முறை சாராப் பிழை (Non Sampling Error)

முதலாம் வகை, ஒரு மாதிரியின் அடிப்படையில் முழுமைத் தொகுதியின் இயல்புகளை மதிப்பீடு செய்கையில், நிகழும் பிழைகளைக் குறிக்கும். முழுமைக் கணிப்பில் இத்தகைய பிழைகள் நேர வாய்ப்பு இல்லை. மாதிரியின் எண்ணிக்கை அளவு (size) அதிகரிக்க அதிகரிக்க, இந்தப் பிழையானது படிப்படியாகக்

குறையத் துவங்கும். ஒரு குறிப்பிட்ட எல்லைக்குப் பின் இப் பிழை குறையும் அளவானது பொருட்படுத்தத் தக்கதன்று. இதன் விளைவாக, அதிகம் செலவாகாமல், குறைந்த முயற்சியுடன் பிழையினைக் குறைக்க இயலும். குறிப்பிட்ட வரையறைக்குட்பட்ட பிழையுடன் மதிப்பீடுகள் பெறப்படுவது மாதிரி அளவீட்டின் முக்கியப் பயன் ஆகும்.

இதற்கு மாறாக, இரண்டாவது வகைப் பிழையானது புள்ளி விவரங்களைப் பல்வேறு வழிமுறைகளில் பயன்படுத்துகையில் ஏற்படுவதாகும். இரண்டு முறைகளிலும் இத்தகைய பிழை ஏற்பட நிறைய வாய்ப்பு உண்டு. முழுமைக் கணிப்பில் இத்தகைய பிழையானது அதிக அளவில் காணப்படும். மாதிரி முறையில், சிறந்த பயிற்சி பெற்ற ஆய்வாளர்களைப் பயன்படுத்துவதன்மூலம் இந்தப் பிழையானது குறைக்கப்படலாம். மாதிரியின் எண்ணிக்கை அளவு (பருமன்) அதிகரிக்கும்போது, இப் பிழையின் அளவும் அதிகரிக்கும். இத்தகைய பிழைகள் நிகழும் விதத்தினையும், இவற்றைக் கண்டறிந்து கட்டுப்படுத்தக் கையாளப்படும் வழிமுறைகளையும், பிறிதொரு பகுதியில் விரிவாகக் காணலாம்.

மாதிரி முறை சார்ந்த பிழையினை (sampling error), மதிப்பிட்ட மாதிரி அளவைக்கும் (statistic) கருதப்பட்ட முழுமைத் தொகுதியில் உள்ள சுட்டுறுப்பின் (parameter) மதிப்புக்கும் இடையில் உள்ள வேறுபாடாக வரையறை செய்யலாம். உதாரணமாக, x என்ற மாறியின் n-உறுப்புகள் கொண்ட மாதிரியின் அடிப்படையில் கணக்கிடப்பட்ட சராசரி மதிப்பீடு $\bar{x}$ என்றும் அதன் முழுமைத் தொகுதி அளவையானது μ என்றும் கொண்டால்,

மாதிரித் தேர்தற் பிழை $(\bar{x}-\mu)$ ஆகும்.

பொதுவாக, θ என்ற முழுமைத் தொகுதி அளவையையும், θ' என்ற அதன் மாதிரி மதிப்பீட்டு அளவையையும் கருதினால், இப் பிழையானது, $(\theta-\theta')$ என்று ஆகும். கொடுக்கப்பட்ட விவரங்களின் இடையில் உள்ள வேறுபாட்டினையும், கையாளப்படும் மாதிரி வழி முறையினையும் பொறுத்து இப் பிழையின் அளவானது மாறும்.

பொதுவாக, x என்ற ஓர் இயல்பின் மாறுபாட்டினைக் (σ) குறைத்து, மாதிரியின் எண்ணிக்கை அளவினை (n) அதிகரிக்க மாதிரிமுறை சார்ந்த பிழையானது மட்டுப்படுத்தப்படும். மாதிரி தெரிவு முறையினைத் தக்க முறையில் மாற்றியமைப்பதன் மூலம்

சராசரிப் பிழையினை மதிப்பிட்டு அறிந்து கட்டுப்படுத்தலாம். நிகழ்தகவினைச் (probability) சார்ந்திருப்பதால், இத்தகைய பிழைகளின் அளவினை மதிப்பிட்டு அறிய இயலும்.

மாதிரிச் செய்முறைகள் தவருகப் பயன்படுத்தப் பெறுவதாலும், விவரங்களைக் குறித்து வைப்பதில் தவறுகள் நிகழ்வதாலும், புள்ளி விவர ஆய்வுக்குக் கருதப்பட்ட நபர்களிடமிருந்து பதிலின்மையாலும் மாதிரிமுறை சாராத பிழைகள் நிகழ்கின்றன. சரியான மாதிரி செய்முறையினைத் தெரிவுசெய்து பயன்படுத்துதல் இன்றியமையாதது. தவருன விவரங்கள், ஆய்வினைத் தவருன வழியில் கொண்டு செல்லும். உதாரணமாக, ஒரு நபரது வருமானம், பெண்ணின் வயது...இவைபோன்ற விவரங்கள் சரிவரக் குறிக்கப்படுவதில்லை. இவை சரியானபடி அமைவது, ஆய்வாளரின் திறனைப் பொறுத்தே அமையும். மாதிரி உறுப்புகளை ஓர் ஒழுங்கற்ற முறையில் தேர்ந்தெடுத்தால், ஒருபுறச் சாய்வான வழு (bias) நிகழ ஏதுவாகிறது. ஆய்வு நடத்தப்படும் மக்களிடம் போதிய அளவு உற்சாகமும் அக்கறையும் இல்லை எனில், விவரங்கள் முழுமை பெறுவதில்லை. எனவே, மனித இயல்புக்கு (human factor) ஏற்ற அடிப்படையில் அமைந்த பிழைகள் இவ்வகையைச் சாரும். இறுதியாக, மாதிரி முறை சாராத பிழைகளை மதிப்பிட இயலாது என்பது குறிப்பிடத் தக்கது.

8. மாதிரிக் கணிப்பும் ஒருபுறச் சாய்வும் (பிறழ்ச்சியும்) (Bias in Sampling)

எடுக்கப்பட்ட ஒரு மாதிரியில் ஒருபுறச் சாய்வுத் தன்மை (அல்லது பிறழ்ச்சி) அமைந்திருப்பின் அம் மாதிரியானது கருதப்பட்ட முழுமைத் தொகுதியின் உண்மையான இயல்பினைப் பிரதிபலிக்க இயலாது. இத்தகைய பிறழ்ச்சியானது பல்வேறு வகைகளில் மாதிரிகளில் வரக்கூடும்.

முதலில், மாதிரி தேர்ந்து எடுக்கும் முறையில் பிறழ்ச்சி ஏற்படலாம். ஆய்வாளரிடம் உள்ள ஒருபுறச் சாய்வு காரணமாக, துணிந்து எடுக்கப்படும் தவருன மாதிரியின் விளைவாக, பிறழ்ச்சி நிகழலாம். மாதிரித் தேர்வில், ஆய்வாளரின் சொந்த விருப்பு வெறுப்புகள் குறுக்கிடுவதால் தவறு நிகழ வாய்ப்பு உண்டு. எவ்வளவு திறமையானவராக இருப்பினும், முழுமைத் தொகுதி பற்றி எவ்வளவு தூரம் தெரிந்திருப்பினும், ஆய்வாளரின் விருப்பத்திற்கேற்ப மாதிரியினைத் தேர்வு செய்ய வாய்ப்பு அளிக்கப்படின், அவ்வாறு தேர்ந்தெடுக்கப்படும் மாதிரியில் பிறழ்ச்சி ஏற்படுவதைத் தவிர்க்க இயலாது. மாதிரித் தேர்வு முறைகள் தீவிர

மாகப் பின்பற்றப்படவில்லை எனில், மாதிரியானது பிறழ்ச்சியி னின்று விடுபட இயலாது. சோதனையாளர் வருகையில், பள்ளி ஆசிரியர் தன் வகுப்பில் உள்ள நன்கு படிக்கும் மாணவனைத் தேர்ந்தெடுத்துக் கேள்வி கேட்பது, ஒருபுறச் சாய்வின் சிறந்த எடுத்துக்காட்டு.

அடுத்து, பிரதியீட்டின் (substitution) விளைவால் ஏற்படும் பிறழ்ச்சி, குறிப்பிட்ட ஒரு நபர் கிடைக்காவிடில், ஆய்வாளர் வேறொரு நபரைப் பட்டியலில் சேர்த்து, அவரைப்பற்றி விவரங் களைக் குறித்தல் இத்தகைய பிறழ்ச்சியை ஏற்படுத்துகிறது. இத் தகைய பிரதியீட்டு முறையானது வாய்ப்பின் அடிப்படையில் (random) அமைய இயலாது. எனவே, ஒருபுறச் சாய்வுத் தன்மையைத் தவிர்க்க முடியாது.

பதிலின்மை (non response) பிறழ்ச்சியின் மற்றொரு காரண மாக அமைகிறது. பதிலளிக்க வேண்டியவர்களில் (respondents) ஒரு சிலர், கேள்விகட்குப் பதிலளிக்காமல் விட்டு விடுகையில், அவர்களுக்கு ஏற்ப மாதிரியின் எண்ணிக்கை அளவைக் குறைக்க இயலாது. பதிலளிக்க மறுக்காத வரை, எவ் வழியிலேனும் தேவைப்படும் விவரங்களை அறிய முற்பட வேண்டும்.

மதிப்பீட்டு முறை (estimation method) மற்றொரு வழியில் பிறழ்ச்சிக்கு வகை செய்கிறது. மாதிரியினின்று கணக்கிடப்பட்ட கூட்டுச் சராசரி (Arithmetic Mean), முழுமைத் தொகுதி சராசரியின் பிறழ்ச்சியற்ற (unbiased) மதிப்பீடாக அமைகிறது. (இதனை விரிவாகப் பின்பு நிரூபிக்கலாம்.) ஆனால், மாதிரியின் மூலம் கணக்கிடப் பெறும் இடைநிலை (Median) மதிப்பானது பிறழ்ச்சியற்ற மதிப்பீடு அன்று. பொதுவாக, $\hat{\theta}$ என்ற மதிப்பீடானது θ என்ற முழுமைத் தொகுதி அளவையின் பிறழ்ச்சியற்ற மதிப்பீடானால் (unbiased estimate),

$$E\left(\hat{\theta}-\theta\right)=0 \text{ அல்லது } E\left(\hat{\theta}\right)=\theta$$

என்பது உண்மையாக வேண்டும். E என்பது எதிர்பார்க்கும் மதிப்பை (expectation) குறிப்பதாகும்.

அதாவது θ என்ற அளவையின் t_n என்ற மதிப்பீட்டு அளவையானது, பிறழ்ச்சியற்றதாக இருக்க வேண்டுமெனில் (n என்பது மாதிரியின் எண்ணிக்கை அளவு), அதன் எதிர் பார்க்கும் மதிப்பானது முழுமைத் தொகுதி அளவை θ வாக இருத்தல் அவசியம். (அ-து.) $E(t_n) = \theta$. மாறி மதிப்புகள் எப்படியும் இருக்கலாம்.

மாறாக, $E(t_n) \neq \theta$ என்று இருப்பின் ($\neq$: சமமில்லை) t_n என்ற மதிப்பீடு θ என்ற முழுமைத் தொகுதி அளவையின் பிறழ்ச்சியுற்ற மதிப்பீடு (Biased Estimator) ஆகிறது. அப் பிறழ்ச்சியானது,

$$B(t_n) = E(t_n) - \theta$$

என்று பெறப்படும். இந்தப் பிறழ்ச்சி மதிப்பு நேர்க்கணிய (+) மானால், t_n என்பது நேர் பிறழ்ச்சி (Positive bias) உடைய தாகவும், எதிர்க்கணிய (—) மானால் எதிர் பிறழ்ச்சி (Negative) bias) உடையதாகவும் கொள்ளப்படும். மற்ற பிழைகள் ஏதும் இல்லை என்று கருதுதல் அவசியம்.

நிகழ்தகவு சார்ந்த பிழையினை மதிப்பிட்டு அறிய இயலும். ஆனால், குறையுடைய மாதிரிக் கூறுகளைத் தெரிவு செய்வதன் மூலமோ அல்லது தவறான விவரங்களைக் குறிப்பதன்மூலமோ அல்லது வேறெந்த விதத்திலோ நிகழும் பிழைகளைத் தொடக்க கட்டத்திலேயே நீக்க முயலுதல் இன்றியமையாதது.

7. இயல்நிலைப் பரவல் (Normal Distribution)

புள்ளியியல் அடிப்படையில், வாய்ப்பு விதிகளுடன் தொடர் புடைய மாறியினைக் (y) கருதுகையில், பெரும்பாலும் மாதிரி ஆய்வுகளிலும் சோதனைகளிலும், பெறப்படும் முடிவுகளிலும் அதிக அளவில் பயன்படுத்தப்படுவது, தொடர்ச்சி (Continuous) தன்மையில் அமையும் இயற்பரவலே ஆகும். நிகழ்தகவினைச் (Probability) சார்ந்து விதிமுறைகளை வகுக்கும் இப் பரவலின் வரைமுறை, y என்னும் தொடர் மாறியினை, (Continuous Variable) பொறுத்த அளவில்,

$$f(y) = \frac{1}{\sigma\sqrt{2\pi}} \cdot e^{-\frac{1}{2\sigma^2}(y-\mu)^2}; \quad -\infty < y < +\infty$$

என்று அமைகிறது. இதன்மூலம், y என்னும் வாய்ப்பு மாறி யானது (Stochastic Variable) (μ, σ^2) என்ற முழுமைத் தொகுதி அளவுகளுடன் இயல்நிலையில் பரவலைக் கொண்டு விளங்கும் என அறியப்படுகிறது. இதனை

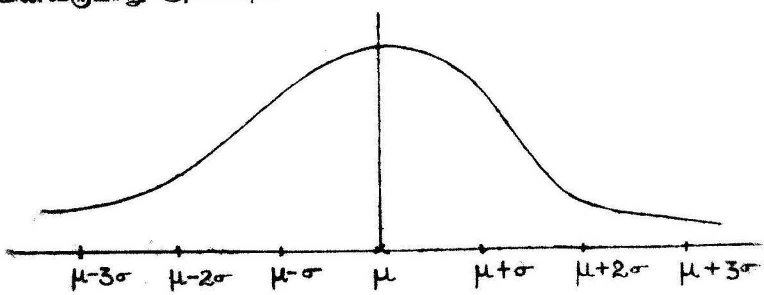
$$y \sim N(\mu, \sigma^2)$$

என்று சுருக்கமாக எழுதுகிறோம்.

இதன் அலைவெண் வளைவானது (Frequency Curve) ஒரே சமச் சீரானதாக அல்லது ஒரே தன்மையானதாக காணப்படும்.

இது ஒரு சமச் சீர் பரவல் (Symmetric Distribution) என்றும் கூறப்படுகிறது.

வரம்புடைய முழுமைத் தொகுதிகள் இறுதியில் இயல்நிலைப் பரவலைக் கொள்ளத்தக்க அளவில் அமைவது குறிப்பிடத் தக்கது. μ, σ^2 மதிப்புகள் கொடுக்கப்படுகையில், இப் பரவலினை நிர்ணயிப்பது எளிதாகிறது. μ (அதாவது மைய மதிப்பு அல்லது சராசரி) மதிப்பினைப் பொறுத்து, இப் பரவலானது சமச்சீர் உள்ளதாக அமைந்திருக்கும். σ-வின் சிறு மதிப்பு களுக்கு, இதன் தன்மை மேல்நோக்கிக் குறுகுவதாகவும் (pealeed) σ-வின் பெரிய மதிப்புகட்குக் கீழ்நோக்கி விரிவதாகவும் (flat) காணலாம். μ-விலிருந்து சில குறிப்பிட்ட தூரங்களில் (Distance) அல்லது தொலைவுகளில் அமையத்தக்க மதிப்புக்கள், பெரும் அளவில் பயன்படுத்தப்படுகின்றன. இயல்நிலை வளைவானது பின்வருமாறு அமையும்:



படம் 1.

μ-வின் இருபுறமும் கருதுகையில், வளைவில் மொத்தப் பரப் புகள் சமமாகப் பரவப்பட்டிருக்கும். y-அச்சில் மாறியின் ஒவ்வொரு y-மதிப்புக்கும் ஏற்ற $f(y)$ மதிப்பு குறிக்கப்படும். μ-விலிருந்து கருதப்படும் σ-தொலைவுகள் (σ-distances) குறிப் பாகக் கருதப்படுகின்றன. உதாரணமாக, $\mu \pm \sigma$ என்ற இடை வெளியில் அமையும் பரப்பினை (Area) இதற்கென அமைக்கப் பட்ட அட்டவணை (table of Normal areas) மூலம் அறியலாம். இது சுமாராக 0·8 என்று அறியப்படும். இது அந்த இடை வெளியில் அமையும் வாய்ப்புநிலை (Probability) யைக் குறிக்கும். இதனை, மொத்த விவரங்களில் சுமார் 80 சதவீதம் $\mu \pm \sigma$ என்ற இடைவெளியில் இடம் பெறும் எனக் கூறலாம். வாய்ப்பு மதிப்பு α என்று ஒரு கெழுவின்மூலம் பொதுவாகத் தரப்பட்டு, வெவ்வேறு மதிப்புகள் விரிவாக அட்டவணைப்படுத்தப்பட் டுள்ளன. [α என்பது நம்பிக்கைக் கெழு (Confidence Coeffi- cients) எனவும் உணரப்படுகிறது.]

பொதுவாக, $\mu \pm K\sigma$ (K வெவ்வேறு நிலைகளில்)-வில். உள்ள வாய்ப்பு நிலைகள் பெறப்படுகின்றன. இவை மதிப்பீட்டுத் திட்பத்திற்கும், நம்பக எல்லை நிர்ணயிக்கவும், முக்கியத்துவ சோதனைக்கும் (Test of Significance) பெரிதும் பயன்படுகின்றன

மாதிரி அளவெடுப்பு முறைகளில் (In Sample Survey Methods)

ஆய்வுகளில் பெரும்பாலும் மாதிரி அளவுகள் பெரிய அளவினதாக (Large Size) அமைய நேரிடுவதால் அவற்றின் மூலம் பெறப்படும் மதிப்பீடுகள் (Estimates) ஏறக்குறைய இயல்நிலைப் பரவலினையே கொண்டு அமைந்திருக்கக் காணலாம். குறிப்பாக, பிறழ்ச்சியற்ற (Unbiased) மதிப்பீடுகளைக் கருதுவோம்.

ஒரு மாதிரிமுறை மூலம் பெறப்பட்ட மதிப்பீட்டை $\hat{\mu}$ என்று கொண்டு, முழுமைத் தொகுதி மதிப்பை μ எனவும் எடுத்துக் கொண்டு, எடுக்கப்படும் எல்லா மாதிரிகளிலும் இம் மதிப்பீட்டினைக் கருதும் நிலையில்,

$$E\left(\hat{\mu}\right) = \mu$$

எனில், $\hat{\mu}$ ஆனது μ-வின் பிறழ்ச்சியற்ற மதிப்பீடாக அமைகிறது. இத்தகைய மதிப்பீட்டைத் தரத்தக்க ஒரு வழிமுறையினைக் கையாண்டு மாதிரியினைத் தெரிவு செல்வதாகக் கொள்வோம்.

மாதிரி மதிப்பீட்டை $\hat{\mu}$ என்றும், இம் மதிப்பீட்டின் மாதிரி முறை சார்ந்த மாறுபாட்டினை σ எனவும் $\hat{\mu}\left[V(\hat{\mu})\right]$ குறிப்பிடுவோம். இந்த மதிப்பீட்டினது திட்பத்தினை (Precision) அறிய முயலுகிறோம். $(\hat{\mu} - \mu)$ என்ற வேறுபாட்டின் மதிப்பினை அறிந்திடாத நிலையில், இயல்நிலைப் பரவலின் சில இயல்புகளை, வாய்ப்பு (Probability) அடிப்படையில் கருதுவது பெரிதும் உதவுகின்றது. இயல்நிலைப் பரவல் இயல்புப்படி, பின்வரும் வாய்ப்பு நிலைகள் அறியப்படுகின்றன:

தனிப் பிழை (Absolute error) என்று கூறப்படும் $|\hat{\mu} - \mu$

ஆனது, $\sigma\hat{\mu}$-ஐ விட அதிகமாக இருந்திட வாய்ப்பு (Chance) 0·32 எனவும்,

$1.96\sigma\hat{\mu}$ (அல்லது, $2\sigma\hat{\mu}$)- வைவிட அதிகமாக இருக்க வாய்ப்பானது 0·05 எனவும்,

மற்றும்

$2{\cdot}58\sigma\hat{\mu}$-வைவிட அதிகமாக இருந்திட வாய்ப்பு 0·01 எனவும், மூன்று குறிப்பிட்ட நிலைகள் அதிகமாகக் கருதப்படுகின்றன.

இவற்றின்மூலம், முழுமைத் தொகுதியின் இயல்புக்கு, கொடுக்கப்படும் நிலைக்கு ஏற்ப, நம்பக எல்லை அல்லது நம்பிக்கை எல்லையினை (Confidence interval) தீர்மானிப்பது எளிதாகிறது. 99% நம்பகத்தன்மை என்பது, ஒவ்வொரு மாதிரியிலும் நம்பிக்கை முடிவுகள் (Confidence statements) பெறப்படும் நிலையில், அவற்றுள் 99% முடிவுகள் சரியாகவும், 1% தவறாகவும் அமைய நேரிடுவதைக் காட்டுவதாகும்.

8. பிறழ்ச்சி மற்றும் அதன் விளைவுகள் (Bias and its effects)

பிறழ்ச்சியின் தன்மையை ஏற்கெனவே கண்டறிந்தோம். மாதிரி ஆய்வுபற்றிய நியதிகளில், பிறழ்ச்சியுடைய மதிப்பீடுகளைச் சில நிலைகளில் கருத வேண்டிய நிர்ப்பந்தம் ஏற்படுகிறது.

இரு முக்கிய காரணங்கள் :

(i) பொதுவான சில பிரச்சினைகளில், அதாவது விகித மதிப்பீடு அல்லது வேறு சில எளிதான மற்றும் பொருத்தமான மதிப்பீடு இவற்றைக் கருதுகையில், பிறழ்ச்சியுடைய மதிப்பீடுகளே பெறப்படும்.

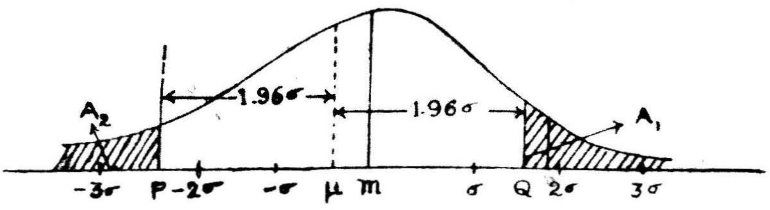
(ii) வாய்ப்பு மாதிரி முறைகளில் பெறப்படும் மதிப்பீடுகள் பிறழ்ச்சியற்றனவாக இருப்பினும், அளவுகளில் ஏற்படும் தவறுகள் மூலமும் (Measurement error), பதிலின்மை மூலம் ஏற்படும் விளைவுகளாலும், பெறப்பட்ட விவரங்கள் அடிப்படையில் நாம் பெறும் மதிப்புகள் பிறழ்ச்சி உள்ளனவாக இருக்கும். உதாரண

மாக, ஏதேனும் ஆய்வினில் பேட்டி கொடுக்க (Interview) மறுப்பு தெரிவிக்கும் நபர்கள், கருதப்பட்ட பிரச்சினைக்கு எதிரான கருத்துகளைக் கொண்டு விளங்குவர். நேரடி பேட்டியில் கருத்துத் தெரிவிப்பவர்கள், சார்பாகவும் எதிராகவும் சம அளவில் பிரிந்து காணப்படுவர். பொதுவாக, எதிரான கருத்துக்கொண்டவர்கள் உண்மையான பதில்களைத் தருவதில்லை; ஆதலால், பிறழ்ச்சி உருவாகக் காரணம் ஏற்படுகிறது.

பிறழ்ச்சியின் விளைவைப் பின்வருமாறு ஆராயலாம்:

சராசரி மதிப்பீடான (Estimate of Mean) $\hat{\mu}$ என்பது 'm' என்ற சராசரி கொண்ட இயல்நிலைப் பரவலை உடையதெனக் கொள்வோம். பிறழ்ச்சியானது

$B = m - \mu$ எனக் கொள்க. (அதாவது m-மதிப்பு, உண்மையான, முழுமைத் தொகுதி சராசரி மதிப்பான 'μ' விலிருந்து மாறுபட்டது. இதனைப் பின்வரும் விளக்கப்படத்தில் கருதுவோம்:



படம் 2.

எந்தவிதப் பிறழ்ச்சி யிருப்பதாகவும் அறியப்படவில்லை என்று கொள்வோம்.

மதிப்பீட்டின் அலைவெண் பரவலைக் கருதி, தரவிலக்கமான σ-அளவினைக் கணக்கிடுவோம். இது m என்ற சராசரியினைப் பொறுத்த அளவில் பெறப்படும் $\sigma_{\hat{\mu}}$ விற்குப் பதில் σ-ஆனது பயன்படுத்தப்படும். மதிப்பீட்டின் துல்லியத்தைப் பொறுத்த

அளவில், $\hat{\mu}$ ஆனது $1 \cdot 96\sigma$ வைவிட அதிக அளவு பிழை உள்ளதாக அமைய நேரிடும் வாய்ப்பு $0 \cdot 05$ என்று கூறுகிறோம்.

இனி, பிறழ்ச்சியின் விளைவால் இவ்வாய்ப்பானது எந்த அளவில் பாதிக்கப்படும் எனக் காண்போம். இதன் பொருட்டு

பிழையானது உண்மையான சராசரியான μ விலிருந்து அளவிடப் படுகையில், மதிப்பீடானது $1{\cdot}96\sigma$ வை விட அதிக அளவு பிழை உள்ளதாக அமைந்திட நேரிடும். ' உண்மையான ' வாய்ப்பின் மதிப்பைக் கணக்கிடல் வேண்டும். கருதப்பட்ட பரவலின் இரு நுனிப் பகுதிகளையும் தனித்தனியாகக் கருதுவோம். மேல்நுனிப் பகுதியில் (Upper Tail Area), பிழையானது $1{\cdot}96\sigma$ வைவிட அதிகமாக இருக்க நேரிடும் வாய்ப்பானது, Q விற்கு மேல் உள்ள (கோடிட்ட) பகுதியாகப்படத்தில் காட்டப்படுகிறது.

வரையறைமூலம், இப் பரப்பானது, (A_1),

$$A_1 = \frac{1}{\sigma\sqrt{2\pi}} \int_{\mu + 1\cdot96\sigma}^{\infty} e^{-\frac{1}{2}\left(\frac{\hat{\mu} - m}{\sigma}\right)^2} d\hat{\mu}$$

என்று அமையும்.

இதில் $\dfrac{\hat{\mu} - m}{\sigma} = t$ என்று பிரதியிட,

தொகையிடற் (integration) பரப்பின் கீழ் எல்லை (Lower Limit) (t-மாறியினைப் பொறுத்து) $\dfrac{(\mu + 1\cdot96\sigma) - m}{\sigma}$

அல்லது $\dfrac{\mu - m}{\sigma} + 1{\cdot}96$

அல்லது $\dfrac{-(m - \mu)}{\sigma} + 1{\cdot}96$

அல்லது $1{\cdot}96 - \dfrac{B}{\sigma}$ [$\because$ $B = (m - \mu)$]

என்று அமைந்திருக்கும்.

$\therefore$ பரப்பானது

$$A_1 = \frac{1}{\sqrt{2\pi}} \int_{1\cdot96 - \left(\frac{B}{\sigma}\right)}^{\infty} e^{-\frac{1}{2}t^2} dt$$

என்பதாகக் காணப்படும்.

இதேபோல், கீழ்நிலை நுனிப் பகுதியை (Lower Tail) கருது கையில், அதன் பரப்பு

$$A_2 = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{-1.9-\left(\frac{B}{\sigma}\right)} e^{-\frac{1}{2}t^2}\, dt$$

என்பதாக அறியப்படும்.

இவற்றிலிருந்து, ஏற்படும் மாற்றம் அல்லது விளைவானது, பிறழ்ச்சி மற்றும் தரவிலக்கம் இவை இரண்டிற்குமிடையில் உள்ள விகித அளவினைப் பொறுத்து அமைவதாக முடிவு செய்யப்படும்.

வெவ்வேறு வாய்ப்புகள், $\left(\frac{B}{\sigma}\right)$ மதிப்புகளுக்கு (0·02 முதல் 1·50 வரை) அட்டவணையிடப்பட்டுக் கொடுக்கப்பட்டுள்ளன. [பார்க்க : W.G. Cochran — 'Sampling Techniques']. 1·96σ அளவினுக்கும் அதிகமான பிழை ஏற்படுவதற்குரிய வாய்ப்புநிலையில், பிறழ்ச்சியானது (Bias) தரவிலக்கத்தின் பத்தில் ஒரு பகுதிக்கு (one tenth) குறைவாக இருப்பின் அதிக அளவில் பாதிப்பு ஏற்படுவதில்லை. பிறழ்ச்சி அதிகரிக்க அதிகரிக்க, வேறுபாடும் (Disturbance) அதிகமாகி, B = σ என்னும் நிலையில் பிழை ஏற்படும் வாய்ப்பு 0·17 என்றாகும். இது முதலில் கருதப்பட்ட மதிப்பைவிட 3 மடங்குக்குமேல் ஆகிறது.

9. மாதிரி அளவெடுப்புகளில் மாறுபாட்டுச் சார்பும், செலவுச் சார்பும் (Variance and Cost Functions)

ஒரு மாதிரி முறை அளவெடுப்புகளை வெற்றிகரமாகத் திட்டமிட்டு நடத்தி முடிப்பதானது, பெரும் அளவில் முழுமைத் தொகுதி மாறுபாட்டினைப் பற்றிய விவரங்களையும், ஆய்வின் வெவ்வேறு நிலைகளில் ஏற்படக்கூடிய செலவுகள்பற்றிய விவரங்களையும் பொறுத்ததாகும். ஆய்வு முடிவுகளில் ஏற்படும் பிழையினை மதிப்பிட, சராசரி வர்க்கப் பிழையினைப் (Mean Square Error) பொதுவாகக் கருதலாம். மதிப்பீடுகள் பிறழ்ச்சி அற்றனவாக இருப்பின், இப் பிழையானது மதிப்பீட்டின் மாதிரி சார்ந்த மாறுபாடாக (Sampling Variance) அமைகின்றன.

மாதிரி அளவானது (Sample Size) அதிகரிக்கையில், மதிப்பீட்டின் மாறுபாட்டு அளவு (Variance of Estimate) குறையவும், அதே சமயம் ஆய்வுச் செலவு அதிகரிக்கவும் செய்கின்றன.

மாதிரிக் கூறுகளைப் பொறுத்தும் இவ்விரண்டும் அமைகின்றன. எனவே, கருதப்படும் ஒரு மாதிரித் திட்டவரையில் (Sample Design), இந்த இரு இயல்புகளையுமே கருத்திற் கொள்ளவேண்டியிருக்கிறது, இதற்குத் தக்கபடியான உத்தம (Optimum) மாதிரி அளவினைத் தீர்மானிக்க வேண்டியது முக்கியக் கடமையாகும். பயிர் சாகுபடிபற்றிய அளவெடுப்புகளில் (Crop Surveys) இவ்விரண்டு இயல்புகளையும் விரிவான முறையில் மஹலநோபிஸ் (Mahalanobis) அவர்கள் கருதியிருக்கிறார்கள்.

மதிப்பீட்டின் மாறுபாடானது, முழுமைத் தொகுதியில் அமைந்துள்ள அடிப்படைக் கூறுகளின் பல்லினத் தன்மை (Heterogeneity), மற்றும் மாதிரியில் இடம் பெறும் கூறுகளின் தன்மை, மாதிரித் தேர்வில் கையாளப்பட்ட வாய்ப்பு அடிப்படை (Probability Scheme), மாதிரி அளவு போன்ற இயல்புகளைச் சார்ந்ததாகும். இச் சார்பினை

$$V(t_n) = f(X, x, p, n,)$$

[t_n என்பது n அளவுள்ள மாதிரிமூலம் பெற்ற மதிப்பீடு.]

இதில் X : முழுமைத் தொகுதிக் கூறின் இயல்பு ;

x : மாதிரிக் கூறின் தன்மை ; p = வாய்ப்பு நியதி ;

n : மாதிரி அளவு என்று அமையும்.

இதனை மாறுபாட்டுச் சார்பு என்கிறோம்.

இதேபோல, ஆய்வுச் செலவினை, மாதிரி அளவு (n), மாதிரிக்கூறின் தன்மை (x), மாதிரி எடுத்த முறை (p) ஆய்வு நோக்கம் (s) ஆகியவற்றினைச் சார்ந்து, $C = \phi(n, x, p, s)$ என்ற செலவுச் சார்பாக (Cost Function) வரையறுக்கப்படும்.

எளிய முறையில் $C = C_0 + nC_1$ என்ற சார்பு பயன்படுத்தப்படுகிறது (C_0 = மாறாச் செலவு ; C_1 : ஆய்வு நடத்த சராசரியாக ஒரு கூறின் செலவு.)

இவை இரண்டும், ஒவ்வொரு வகையான மாதிரித் தேர்வு முறைபற்றிய விரிவான ஆய்விலும் முக்கிய இடம் அளிக்கப்பட்டுள்ளன.

எனவே, சுருக்கமாக, கொடுக்கப்படும் திட்ப அளவில் (Precision), குறைந்த செலவுடன் அமையும் முறையானது பெரிதும் விரும்பப்படும்.

10. மதிப்பீட்டுத் திட்பம் (Precision)

மாதிரிமுறை பின்பற்றப் படுவதன்மூலம் பெறப்படும் மதிப்பீடுகள் (Estimates) எவ்வளவு தூரம் அதே சூழ்நிலையில் பெறப்பட்ட முழுமைத் தொகுதி மதிப்புகளிலிருந்து விலகியுள்ளன என்பதை அறிய, ஒவ்வொரு மதிப்பீட்டிற்கும், மாதிரி முறை சார்ந்த மாறுபாடு (Sampling Variance) கணக்கிடப்படுகிறது. அதாவது மாதிரி அளவெடுப்பு மேற்கொள்ளப்பட்ட அதே சூழ்நிலைகளில், அதே அளவு முறைகள், அதே ஆய்வுமுறை, அதே ஆய்வாளர் போன்றவற்றைக் கருதி, முழுக்கணிப்பு மேற்கொள்ளப்பட்டு முடிவுகள் பெறப்படின், அவை எவ்வளவு தூரம் மாதிரி மதிப்பீடுகளிலிருந்து விலகியிருக்கும் என்று காண்கிறோம். இத்தகைய வேறுபாடானது மதிப்பீட்டு திட்பம் (Precision of Estimate) அல்லது மாதிரி மதிப்பீட்டின் நம்பகத்தன்மை (Reliability) என்று கூறப்படும். மதிப்பீட்டுத் துல்லியமானது (Accuracy) மாதிரி மதிப்பீட்டுக்கும், உண்மையான மதிப்புக்கும் உள்ள வேறுபாடாகும். வாய்ப்பு அடிப்படையில் எடுக்கப்படும் மாதிரிகளில், அம்மாதிரிகள் மூலமாகவே திட்ப அளவானது முடிவு செய்யப்படும். மாதிரிமுறை சார்ந்த பிழைகட்கு (Sampling errors) வெளி எல்லைகளை (Outer bound) வரையறுத்து, இவற்றுக்கு அப்பால் செல்லக்கூடிய (exceeding) வாய்ப்பானது மிகக் குறைவானதாக ஆக்கப்படுகிறது. வேறொரு பகுதியில், மாதிரி அளவினை (Sample size) நிர்ணயம் செய்யும் பிரச்சினையில் திட்ப அளவு பெறும் இடம் வகிப்பதை விரிவாகக் கருதுவோம்.

11. மாதிரி முறையின் இரு தத்துவ நியதிகள் (Two Principles of Sampling Method)

மாதிரிகளைத் தெரிவு செய்து அவற்றிலிருந்து பெறப்படும் மதிப்பீடுகள், முழுமைத் தொகுதியின் இயல்புகளைத் திறனுடன் பிரதிபலிக்கும் தன்மையானது, பின்வரும் இரு அடிப்படை விதிகள் (laws) மூலம் அமைவதாகும்.

(i) புள்ளிவிவர ஒழுங்குமுறை விதி (Law of Statistical Regularity)

இதன்படி சம வாய்ப்பின் அடிப்படையில் எடுக்கப்பட்ட மாதிரியானது, அது எடுக்கப்பட்ட முழுமைத்தொகுதிஇயல்புகளை கூடிய வரையில் கொண்டு விளங்கும் என்று கூறுகிறது. முறையற்ற அல்லது விதி வழுவிய முறையில் எடுக்கப்பட்ட மாதிரியானது தவறானதாக (Biased) அமைகிறது. ஒரு தொகுதி

யினின்று எடுக்கப்பெறும் சமவாய்ப்பு மாதிரியானது (Random Sample), அத் தொகுதி உறுப்புகட்கு சமஅளவில் தேர்வுபெறும் வாய்ப்பளித்து. இயல்புகளின் பிரதிபலிப்பாய் அமைகிறது.

(ii) **பேரினங்களின் மாறாப் பொதுமை விதி** (Law of Inertia of Large Numbers)

இந்த விதியின்படி, பெரிய எண்ணிக்கை, (Size) அளவில் எடுக்கப் பெறும் மாதிரியானது சிறிய எண்ணிக்கையில் எடுக்கப் பெறும் மாதிரியினைவிட மாறுநிலையினைப் பெற்றதாக இருக்கும். சிறிய எண்ணிக்கையில் எடுக்கப்பட்ட மாதிரி நம்பகத் தன்மை குறைந்து விளங்கும். உதாரணமாக, நூறுபேர்கள் அடங்கிய தொகுதியின் சராசரி வருமானத்தைக் கணக்கிடுதற்கு, இரண்டு அல்லது மூன்றுபேரை மட்டும் தேர்ந்தெடுத்து, சராசரிவருமானத் தைக் கணக்கிடுவது சரியான முறையாக அமையாது. எடுக்கப் படும் மாதிரியின் எண்ணிக்கை முப்பதாக அதிகரிக்கப் படுகையில், மாதிரியின் சராசரி முழுமைத் தொகுதி மதிப்புக்கு ஏறக்குறைய சமமானதாக அமைவதைக் காணலாம். எனவே அடையப்பெறும் முடிவுகள், மாறுநிலைபெற்றும், பொருத்தமான வையாகவும், திட்பமானவையாகவும் அமைவது, எடுக்கப்பட்ட மாதரியின் எண்ணிக்கையைப் பொறுத்ததாகும்.

மேற்கூறப்பட்ட இருவிதிகளின்படி, சிறந்த பயனுள்ள, நம்பகம் வாய்ந்த, நடைமுறைக்கு ஏற்றமாதிரி எடுத்தல் முறை யானது சமவாய்ப்பின் அடிப்படையிலும், எண்ணிக்கைஅளவில் அதிகமானதாகவும் அமைந்திருக்க வேண்டும். எடுக்கப்படும் மாதிரியின் தன்மையைப் பொறுத்தே, மற்றக் கணிப்பு முறைகள் அமையும் ஆதலால், எடுத்துக் கொண்ட நோக்கத்திற்கு ஏற்ற வகையில் மாதிரிக் கணிப்பு முறையினைப் பின் பற்றுதல் இன்றி யமையாததாகிறது.

3. மாதிரி அளவெடுப்பினைத் திட்டமிடுதல்

(Planning of Sample Survey)

[அளவெடுப்பின் நோக்கம் (குறிக்கோள்) — புள்ளி விவரங்கள் — புள்ளிவிவரங்களை சேகரித்தல் — வினாப்பட்டியல் — காலவரைகள் — மாதிரிச் சட்ட அமைப்பு — மாதிரி முறைத் திட்ட அமைப்பு — முன்னோடி ஆய்வு — களப்பணி — விவரங்களின் பகுப்பாய்வு — செயல்முறைப் படுத்தல் — இறுதி அறிக்கை — மாதிரி அளவு நிர்ணயித்தல் — குறியீடு பற்றிய விளக்கம்.]

அளவெடுப்பின் நோக்கம் (Scope)

புள்ளியியல் ஆய்வானது, பொதுவாகப் பல்வேறு புள்ளியியல்முறைகளைப்பின்பற்றி, ஒரு குறிப்பிட்ட பிரச்சினை (அல்லது பொருள்) பற்றி நன்கு ஆராய்ந்து கண்டறியும் முடிவுகளைப் பற்றியதாகும். முழுக் கணிப்பில் உள்ள பல்வேறு கட்டங்கள் (Stages), மாதிரி அளவெடுப்பு முறைக்கும் பொருந்தும். பல்வேறு நிலைகளை அதிக சிக்கல்கள் இல்லாது கடந்தபின்னரே, முடிவுகள் பெறப்படுகின்றன. ஒவ்வொரு கட்டத்திலும் ஒரு அடிப்படைத் தேவை பூர்த்தி செய்யப்படுதல் வேண்டும். அதிக கவனமும் தேவைப்படுகிறது. அளவெடுப்பினைத் திட்டமிடும் நிலையிலேயே சில விவரங்களை நன்கு கருத்தில் கொள்ள வேண்டும்.

1. அளவெடுப்பின் அடிப்படை, நோக்கம்பற்றிய விரிவான அறிக்கை.
2 புள்ளி விவரத் தேவைகள்பற்றிய வரையறுப்பு.
3. புள்ளி விவரங்களைக் கண்டறியும் முறைகள்
4. வினாப்பட்டியல் அமைப்பு (Questionnaire)
5. மேற்கோள்கள் (Reference) மற்றும் அறிவிக்கும் கால அளவு.
6. மாதிரி முறைக்கான ‘சட்ட’ வரையறை (Frame)
7. மாதிரியின் திட்ட அமைப்பு (Design) தெரிவு செய்தல்
8. முன்னோடியான கணிப்பின் (Pilot Survey) திட்டம்
9. கள ஆய்வு (Field Work).
10. பலநிலை கடந்து, இறுதி அறிக்கை தயாரித்தல்.

மேற் கூறப்பட்ட விவரங்கள் அனைத்தும் ஒன்றோடு ஒன்று மிகவும் தொடர்புடையனவாகும். இவற்றைப்பற்றி எடுக்கப்படும் முடிவுகள் இணைந்தனவாக அமைய வேண்டும். வெவ்வேறு விவரம் குறித்து எடுக்கப்படும் முடிவுகள் தகுந்த கட்டங்களில் அல்லது நிலைகளில் சரியானபடி இணைக்கப்படவேண்டும். மேற் குறிப்பிட்ட பல்வேறு விவரங்களையும் மிக விரிவான அளவில் எடுத்து ஆராயலாம். சில கருத்துக்கள் மிகவும் முக்கியமானவை யாக அமைகின்றன.

அளவெடுப்பின் நோக்கம் (Objective of the Survey)

முதன் முதலாக, கருதப்படும் பிரச்சினை அல்லது பொருள் பற்றித் தெளிவாக வரையறுக்கவேண்டும். இதன்மூலம், கணிப்பில் இடம்பெறும் பல்வேறு காரணிகளைப்பற்றி ஆய்வாள ருக்குத் தெளிவாகப் புலனாகும். ஒரு முடிவினை அடையும் முன், எந்தவித முக்கிய விவரமும் விட்டுப்போகாவண்ணம் இது உதவு கிறது. தேவையற்ற விவரங்களை அடியோடு களைந்திடவும் இது வழி செய்யும். கால சிக்கனமும் பொருள் சிக்கனமும் இதனால் பெறப்படும். ஆய்வினை நடத்த எடுத்துக் கொள்ளப்படும் பூகோளப் பரப்பைத் தெள்ளத்தெளிய முன்கூட்டியே அறிவிப்பது இன்றியமையாதது. ஒரு நிலப்பரப்பில் கிடைக்கும் பலவகை விவரங்களைத் தீவிரமாகக் கண்டறியமுடிகிறது. ஒரு குறிப்பிட்ட கால அளவில் ஆய்வினை முடிக்கவேண்டும் என்ற நிபந்தனையும் விதிக்கப்படவேண்டும். கவரப்படும் பரப்பினைப் பொறுத்தே, ஆய்வின் மற்றெல்லா நிலைகளும் அமையும் என்பது குறிப்பிடத் தக்கது. என்ன காரணத்துக்காக ஆய்வு (Survey) நடத்தப்படு கிறது என்பதே முக்கியம். இதைத் திட்டவட்டமாகக் குறிப்பிட் டால்தான், எந்தக் கணிப்பு முறையும் 100% வெற்றியடையும். நிறுவப்படும் உண்மைகள் பயன்படுத்துவோரின் கண்ணோட்டத் தில் ஏற்கக்கூடியனவாக அமையவேண்டும்.

2. புள்ளி விவரங்கள் பற்றிய தெளிவு (Data Requirements)

நிகழ்காலத்திற்கும் எதிர்காலத்துக்கும் பயன்படக்கூடிய வகையில் புள்ளி விவரங்கள் சேகரிக்கப்படல்வேண்டும். தேவை யான விவரங்கள் பற்றிய தெளிவு, ஆய்வுப்பரப்பு, விவரங்களின் அட்டவணைப்படுத்தும் முறை, எதிர்பார்க்கும் துல்லிய மதிப்பு, ஆய்வுச் செலவு (cost) இவைபற்றி நன்கு அறிந்திருத்தல் அவ சியம். நாட்டு நலத்துக்கும் மக்கள் நலத்துக்கும் வகுக்கப்படும் திட்டமானது புள்ளி விவரத் தன்மையினைப் பொறுத்து நன்மை தரும். எனவே மக்கட்தொகை, பிறப்பு-இறப்பு வீதம், வரவு-

செலவு மதிப்பு, சேமிப்பு-கடன் பற்றிய கணக்கீடு, கல்வி நிலை, விவசாய வளர்ச்சி, விலைவாசி நிலைமை போன்ற பல்வேறு உண்மைகள் பற்றிய விவரங்கள் அனைத்தையும் அறியவேண்டியுள்ளது.

புள்ளி விவரத் தேவைகளை வழிமுறைப்படுத்தும்போது, மதிப்பிட்டு அறியவேண்டிய முழுமைத் தொகுதிக் கூட்டு மதிப்புகள், சராசரிகள், விகிதங்கள் போன்றவற்றையும், இவற்றைக் கணிக்கவேண்டிய கூறுகளின் தொகுப்பினையும், மதிப்பீட்டின் கால அளவினையும் (மாத அடிப்படையில், ஆண்டு அடிப்படையில்... என்றவாறாக) குறிப்பிட்டுக் காட்டவேண்டும். மதிப்பீடுகள் பெறவேண்டிய பிரிவுகளையும் (Classification), நிலப்பரப்பையும் (Geographical Area) தெளிவாகக் காட்டவேண்டும். தேவைப் பூர்த்தி அடிப்படையில் விவரங்கள் பெறப்படவேண்டும். முழுமைத் தொகுதியை நன்கு நிர்ணயம் செய்ய வேண்டும்.

மேலும், மதிப்பீடு செய்யும்போது, ஏற்படக்கூடிய பிழையின் எல்லை அளவை (Margin of Error) உறுதி செய்தல் அவசியம். 'அனுமதிக்கக்கூடிய பிழை' (Permissible Errors) எந்த வித இயல்புக் கணிப்பிலும் வரக்கூடியன ஆகும். கணிப்பு ஆய்வின் அடிப்படையில் பெற்ற முடிவுகளின்பேரில் எடுக்கும் நடவடிக்கைகளைப் பொறுத்து இத்தகைய பிழைகள் உறுதிப்படுத்தப்படுகின்றன. ஆனால் செயல்முறை சிக்கல்கள் காரணமாக, இது எளிதாக அமைவதில்லை. முன்பே (Past) நடைபெற்ற ஆய்வில் பெற்ற முடிவுகளை வைத்து, இப்பிழைகளைக் கட்டுப்படுப்படுத்த முயலலாம். நடைமுறையில், கணிப்பு ஆய்வின் செலவானது (Cost of the Survey) கட்டுப்படுத்தப்படுவதால், புள்ளி விவர ஆய்வாளரின் திறனைப் பொறுத்து, திட்பமான (Precise) மதிப்பீடுகள் பெறப்படுகின்றன.

அளவெடுப்பின் குறிக்கோளுக்குத் துணைசெய்யும் சில அதிகப்படி விவரங்கள் (Additional Informations) சேர்க்க வாய்ப்பு உண்டு. பெரிய அளவில் (large scale) நடத்தப்பெறும் பகுப்பாய்வில் இது மிகவும் உதவுகிறது.

தற்போதைய நிலையைமட்டும் அறிய மேற்கொள்ளப்படும் 'அட்-ஹாக்' (Ad-hoc) அல்லது தற்காலிக ஆய்வு முறையில் சில நிபந்தனைகளைத் தீவிரமாகப் பின்பற்றத் தேவையில்லை. ஆனால் அடிக்கடி நடத்தப்பெறும் ஆய்வுகளில், காலத்தோடு இயைந்த

போக்குபற்றி (Trend) சில இயல்புகளை ஆராயும்போது, தொடர்ந்து ஆய்வு நடத்தப்படுவதால், கணிப்பு முறையினை மிகச்சிறந்த முறையில் திட்டமிடல் இன்றி யமையாததாகும். மேலும், அதிகத் திறமை வாய்ந்த ஆய்வாளரைப் பயன்படுத்த இயலுவதால், செலவினைக் கட்டுப்படுத்த வாய்ப்பு உண்டு. எதிர் காலக் கணிப்புகளில் தகுந்த மாற்றங்கள் செய்ய வழி அமை கிறது. முதலிலேயே தக்க கவனம் செலுத்தி விவரங்கள் சேக ரிக்கப்படின், பின்நிலைகளில் பிழையற்ற, துல்லியமான (Accurate) முடிவுகளைப் பெற இயலும்.

3. புள்ளி விவரங்களைச் சேகரிக்கும் முறை (Data Collection Method)

ஆய்வுக்குத் தேவைப்படும் புள்ளி விவரங்களைப் பல முறை களில் கண்டறியலாம்:

(i) பருப்பொருள் அடிப்படையில் அளந்து கண்டறிதல் (Physical Measurement)

(ii) நேரடி பேட்டியின் (தொடர்பு) விவரம் சேகரித்தல் (Direct Interview)

(iii) அஞ்சல் வழி ஆய்வு முறை (Mail Enquiry)

(iv) பதிவு முறை (Registration)

(v) பழைய குறிப்பேடுகளிலிருந்து விவரங்களை எடுத்தல் (Transcription from Records)

முதல் நான்கு முறைகளின்மூலம் முதல் நிலை விவரங்கள் (Primary Data) பெறப்படுகின்றன. இவை முதன்முறையாக, பதிலளிப்போர் வாயிலாக (Respondents), அல்லது அலகுகளி லிருந்து, சேகரித்து அறியப்படும் விவரங்களாகும். இவற்றில் கையாளப்படும் முறைகள் இங்கு விரிவாக எடுத்துரைக்கப் படுகின்றன.

இறுதியாக (v)-ல் உரைக்கப்பட்ட வழிமுறை மூலமாக எடுக் கப்படும் விவரங்களை, இரண்டாம் நிலை விவரங்கள் அல்லது வழி நிலை விவரங்கள் (Secondary Data) என்று வரையறுக்கலாம்.

எந்த ஒரு கணிப்பிலும், அதன் தன்மைக்குப் பொருத்தமான இயைவான ஒரு முறையினையே பின்பற்றி, விவரங்கள் கண்டறி யப்படுதல் வேண்டும். கணிப்பாய்வின் (Survey) திட்டமிடலும், நல்ல முறையில் நிறைவேற்றி முடிப்பதும், விவரங்களை சேகரித்

துக் கண்டறியக் கையாளும் முறையினைப் பொறுத்தே அமைந்திடக் காணலாம். பெறப்படும் துல்லியம் (Accuracy) நடைமுறை இயைவு (Practicability), மேலும் ஆய்வுச் செலவு (Cost) இத்தகைய தன்மைகளைக் கருத்திற் கொண்டு, சிறந்த முறையினைத் தேர்ந்து எடுத்து விவரங்களைக் கண்டறிதல் மிக அவசியமாகிறது. விவரச் சேகரிப்பு முறை, மாதிரிக் கணிப்பு மற்றும் முழுமைக் கணிப்பு ஆய்வுகளில் எழும் முக்கியப் பிரச்சினையாகும். பல்வேறு முறைகள்பற்றி ஒரு விரிவான விளக்கத்தினை, மேல் வரும் பத்திகளில் காணலாம்.

(i) அளந்து கண்டறியும் முறை (Physical Measurement)

ஒரு அலகின் அல்லது கூறின் தன்மையினை அல்லது பதில் அளிப்போரின் இயல்பை, ஒரு அளவீட்டுக் கருவியின் துணை கொண்டு (Measuring Instrument), (பரப்பொருள் அடிப்படையில்—Physically) அளவிடும் முறை இது. உதாரணமாக, உடற்கூற்று இயல் (Anthropology) ஆய்வுகளில், ஒரு மாதிரிக் கூறில் தேர்வு பெற்ற நபர்களின் தலை-அகல (குறுக்கு), அளவு (head-breadth), கைகளின் நீட்டளவு (length of hands) போன்றவற்றை அளந்து கண்டறியலாம். விவசாயத் துறை ஆய்வில் (Agricultural Survey), கொடுக்கப்பட்ட நிலப்பரப்பில் கருதப்பட்ட பயிரானது விளைவிக்கப்படும் பகுதி அளவினை, ஒவ்வொரு மாதிரிக்கூற்றுக்கும் நிர்ணயிக்கலாம். இதுகண்கூட்டுக் கண்டறிவு முறை (Eye-observation). இம்முறையைக் கடைபிடிக்கையில், ஆய்வுச் செலவினையும், விவரங்களின் இயல்பினையும் கருத்தில் கொள்ளுதல் முக்கியம். எனினும், சில துறைகளில் இம்முறை அத்தியாவசியமாகத் தேவைப்படும். எடுத்துக்காட்டாக ஒரு குறிப்பிட்ட நோயின் நடமாட்டத்தை அறிய, தக்க மருத்துவப் பரிசோதனை தேவைப்படும் நிலையில் இம்முறை பயன்படுத்த வேண்டிய ஒன்று.

(ii) நேரடிப் பேட்டி முறை (Personal Interview)

பதிலளிப்போரை நேரடியாக அணுகி, வினாக்களை எழுப்பி அவற்றின் மூலம் விவரங்களை சேகரிப்பது நேரடித் தொடர்பு முறை (Direct Method). சமூக மற்றும் பொருளாதார ஆய்வு முறைகளில் இந்த முறை மிக அதிக அளவில் பயன்படுத்தப்படுகிறது. நேரடியாகக் காணுவதன்மூலம், பதிலளிப்போரிடமிருந்து கூடுமானவரை பிழையற்ற அல்லது சரியான (accurate) விவரங்களைப் பெற வாய்ப்பு உண்டு. ஆய்வின் நோக்கத்தினைத் தெளிவாக எடுத்துரைத்து உணர்த்துவதன்மூலம், தேவையை வலியுறுத்து

வதன்மூலம், தேவைப்படும் விவரங்களைத் திரட்ட முடிகிறது. இதில், 'பதிலின்மை'யின் அளவு (Non-response) மிக மட்டுப் படுகிறது. பெறுதற்கு மிக அரிய சில விவரங்களையும், திறமை யின்மூலமாகத் தெரிந்துகொள்ள ஏதுவாகிறது. (உதாரணம்: பெண்களின் வயதுபற்றி; வருமான உண்மைகள்......)

(iii) அஞ்சல் வழி ஆய்வு (Mail Enquiry)

அஞ்சல் மூலம் வினாப்பட்டியல்களைப் பதிலளிப்போருக்கு அனுப்பி, அவர்களிடமிருந்து பெறப்படும் பூர்த்தி செய்த பட்டிய லின்மூலம் விவரங்களைக் கண்டறியும் முறை இது. இம் முறை யில் ஆய்வுச் செலவு கணிசமான அளவு அதிகரிக்கிறது. ஆனால், இம்முறையில், பெரும் அளவில் பதிலின்மை (Non-response) அமைய வாய்ப்பிருப்பதால், ஆய்வுமூலம் பெறப்படும் உண்மை களும் பெரும் அளவில் பாதிக்கப்படுகின்றன. பதிலளிப்ப வர்கள் வினாப்பட்டியலில் அடங்கிய விவரங்களை நன்கு புரிந்து கொண்டு, பதில்களைத் தருவதற்கு பெரும் அளவில் இசைவுதரின், பதிலின்மையின் அளவு குறைந்திடும். இம்முறையுடன் கூடவே, பதில் கிடைக்காத இடங்களில் நேரடித் தொடர்பினைக்கையாண்டு விவரங்களைப் பெறவேண்டியிருக்கும். எளிதில் புரிந்துகொள் ளத்தக்க விவரங்களைப் பற்றி அறிவதற்கு இம்முறை உதவு கிறது. பெரும்பாலும் இம்முறையானது பயன் தருவதில்லை. நடைமுறைச் சிக்கல்கள் அதிக அளவில் காணப்படுகின்றன. மேற்கத்திய நாடுகளில் இம்முறை பெரிதும் பயன்படுத்தப்படு வதாகும்.

(iv) பதிவு முறை (Registration)

இதன்படி, பதிலளிப்போர் (அதாவது ஆய்வுக்கு உட்பட் டோர்), தேவைப்படும் விவரங்களை சில குறிப்பிட்ட இடங்களில் பதிவு செய்து வைத்தல் வேண்டும். பிறப்பு-இறப்புபற்றிய இய லில் (Vital Statistics), இம்முறை பயன்தருகிறது. அதாவது, ஒரு இடத்தில் வாழும் மக்கள், அங்கு நிகழும் ஒவ்வொரு பிறப்பினையும் இறப்பினையும் தக்க அலுவலர்களிடம் பதிவுசெய்யு மாறு கோரப்படுகின்றனர். அஞ்சல் முறை, நேரடித்தொடர்பு முறை ஆகிய இரண்டின் கலப்பு இம்முறையாகும். தேவைப் படும் விவரங்கள் குறித்து பதிலளிப்போர் அறிவிக்கப்படுகின் றனர்; அதன்பின், பதிவுசெய்யப்படும் விவரங்களின் உண்மை நிலையை உறுதி செய்து, குறித்து வைக்கின்றனர். தகுந்த அலு வலர்கள். விருப்பின்மை, கவனக் குறைவு ஆகியவற்றின் காரணமாக, பதிலின்மை அளவுமிகுந்து காணப்படுகிறது. கண்

டிப்புடன் வற்புறுத்தப்படும்போது (compelled) ஓரளவுக்கு பதிலின்மை குறைய வாய்ப்பு உண்டு. மாதிரிக் கணிப்புக்கு, மாதிரிப் பதிவு முறை (Sample Registration) உதவுகிறது.

(Ref : மாதிரிப் பதிவீட்டு முறை)

(iv) **ஏற்கெனவே சேகரிக்கப்பட்ட விவரங்களைப் பயன்படுத்துதல் (Transcription from Records)**

ஏதேனும் குறிப்பிட்ட (special) ஆய்வுக்கு விவரங்கள் தேவைப்படின், ஏற்கெனவே பல்வேறு துறைகளில் சேகரித்துப் பதியப்பட்ட பதிவேடுகளைப் பயன்படுத்தி விவரங்களை அறியலாம். மூலக் கூறுகளைக் கருதி முதன்முறையாக ஆய்வு நடத்த வேண்டிய சிரமம் இந்த முறையில் வருவதில்லை. செலவும், நேரமும் மிச்சப்படுத்தப்படுகின்றன. அரசுத்துறைப் புள்ளி விவரங்கள் (Official Statistics) மற்றும் தொழிற் புள்ளிவிவரங்கள் (Business Statistics) ஆகியவை ஏற்கெனவே சேகரித்து வைக்கப்படுவதால், எந்த மாதிரிக் கணிப்பாய்விலும் அவற்றை அப்படியே எடுத்துக்கொண்டு கருதுவது எளிதாகிறது. உதாரணமாக, மக்கட் தொகைக் கணிப்பில் (Population census) 'ஆய்வுச் சீட்டு' (Enumeration-slip) மூலமாகப் பெறப்படும் பல்வேறு உண்மைகள் (எடுத்துக்காட்டு : பணி புரியும் நபர்கள் மற்றும் வேலையற்றிருப்போர் பற்றிய விவரங்கள்) எந்த ஒரு மாதிரிக் கணிப்பு ஆய்வுக்கும் மிகவும் உதவி புரிகின்றன.

அடுத்து, மிகவும் முக்கியமான ஒரு கருத்தினை எடுத்துக் கொள்வோம். மாதிரிக் கணிப்பின் திட்ட அமைப்பு (Design) இக் கருத்தின் தன்மையைப் பொறுத்தே அமைவதைக் காணலாம்.

4. வினாத்தொகுதி அல்லது வினாப்பட்டியல் (Questionnaire *Vs* Schedule)

இத்தகைய பட்டியலை அமைக்கு முன்னர், இம் முறை கருதப்பட்ட மாதிரி ஆய்வுக்கு (Sample Survey) தேவைதானா என்பதைத் தெளிவாக்கிக் கொள்ளவேண்டும். நம்பத்தக்க அளவில் உண்மை விவரங்களை அறிந்துகொள்ள இப்பட்டியல் தேவைப்படும்.

வினாப்பட்டியல் முறையில், தகவல் அளிப்போர் அல்லது பதிலளிப்போரிடம் (Informants or Respondents), முன் கூட்டியே தீர்மானிக்கப்பட்ட வினாக்கள் அடங்கிய தொகுதி அல்லது பட்டியல் அளிக்கப்பட்டு, அவற்றுக்கேற்ற பதில்கள்

அவர்களாலோ அல்லது ஆய்வாளர்களாலோ குறித்துவைக்கப் படுகின்றன.

இதற்கேற்ப, பதிலளிப்போர் கேட்கப்படும் வினுக்களைத் தெளிவாக, ஐயந்திரிபற்றுப் புரிந்துகொண்டு தக்க பதில்களைத் தருமாறு எதிர்பார்க்கப்படுகின்றனர். புரியாத சொற்கள் அல்லது கருத்துகட்கு (Terms) விளக்கம் கூறும் வகையில், ஆய்வாளர் எந்த வகையிலும் பதிலளிப்போரது சொந்தக் கருத்தினைத் திசை திருப்பலாகாது. இந்த முறையானது, அஞ்சல் வழி ஆய்வில் பெருமளவில் பயன்படுகிறது. ஆய்வு நடைபெறும் முன்னரே வினுப்பட்டியலானது தயார் செய்யப்படுதல் வேண்டும்.

ஆய்வாளரின் பொது அறிவு நிலை, கூர்மதி, சிறந்த அனுபவம் இவற்றைப் பொறுத்து ஒரு சிறந்த வினுப்பட்டியலானது அமைந்திருக்கும். வினுப் பட்டியல் முறைப்படி (Schedule) கேள்விகளை அமைத்து, மதி நுட்பத்துடன் தேவையான பதில் களைப் பெறுவது ஆய்வாளரின் சாமர்த்தியத்தைப் பொறுத்தது. தனது அனுபவ அறிவைப் பயன்படுத்திப் பல்வேறு கருத்துக்களை யும் வரையறைகளையும் நன்கு விளக்கிக் கூறி, பொருத்தமான விவரங்களைப் பெறுதற்கு ஆய்வாளர் முயலவேண்டும். இவற்றைப் பற்றியும், கையாளவேண்டிய முறை பற்றியும், ஆய்வாளருக்கு ஏற்கெனவே விளக்கம் கொடுக்கப்பட்டு இருக்கவேண்டும்.

இந்த வினுப்பட்டியல் ஆய்வு முறையில், ஆய்வாளர்மூலம் நிகழும் பிறழ்ச்சி அல்லது ஒருபுறச் சாய்வுக்கு (Bias) அதிக வாய்ப்பு ஏற்படுகிறது. ஏனெனில் அவர்கள் தங்கள் விருப்பு வெறுப்புக்கு ஏற்றபடி பதிலளிப்போரை மாற்றிவிடமுடியும் [வினுத்தொகுதி (Questionnaire) முறையில், பதிலளிப்போர் தாங் களாகவே விவரங்களைப் பூர்த்திசெய்து தருகின்றனர். கேள்வி களும் முன்கூட்டியே நிர்ணயிக்கப்படுகின்றன.] சிறந்த முறை யில் பயிற்சிபெற்று திறமை பெற்ற ஆய்வாளரைப் பயன்படுத்து வதன்மூலம், இப் பிறழ்ச்சியை ஓரளவு தவிர்க்கலாம். ஆனால், வினுத்தொகுதி முறையில் (Questionnaire), சிக்கலான கேள்விக ளாலும், புரிந்துகொள்ள இயலாத கருத்துக்களாலும், பதிலளிப் போர் சார்பான பிறழ்ச்சி ஏற்படுகிறது. இதற்கு, நன்கு பயிற்சி பெறும் ஆய்வாளர் சுருக்கமாக விளக்கம் தருமாறு அமைதல் வேண்டும். இது சற்றுக் கடினமான காரியமே. தக்க விளக்கங் களுடன் கூறுகள் (Items) அமைந்திருப்பின், தேவையான விவரங்கள் ஆய்வாளருக்குக் கிட்டுகின்றன. வினுத்தொகுதி (Questionnaire) அமைப்பில், பட்டியல் முறையை (Schedule) விடச் செலவானது குறையவழி உண்டு.

மாதிரி ஆய்வினைத் திட்டமிடுகையில் வினாப்பட்டியல் அமைப்பில் (Preparation) தனிக் கவனம் செலுத்தப்படல் வேண்டும். ஆய்வில் பெறப்படும் முடிவுகள் இதன் தன்மையின் அடிப்படையில்தான் அமைந்திருக்கும். வினாக்களை, எளிய முறையில் தெளிவாக, உற்சாகம் தரத்தக்கனவாக, திறமையாக அமைக்கவேண்டும். பதில்களைக் கோடிட்டுக் காட்டும்படியாக கேள்விகள் அமைதல் கூடாது. வினாக்கள் இயல்பான தொடர்புடன் (Sequence) அமையவேண்டும். பொதுவான விவரங்களை முதலில் கருதி, அதன் பின்னர் கருதப்பட்ட ஆய்வின் தன்மைக் கேற்ற வினாக்களை எழுப்ப வேண்டும். எளிதில் அட்டவணைப் படுத்தத் தக்க விதத்தில் கூறுகள் (Items) அமையவேண்டும். கூடிய வரையில் பதில்களை எண்களாக (Numbers) அல்லது குறியீடுகளாக (Codes) குறிக்கும் வண்ணம் வினாக்கள் இருக்க வேண்டும்.

சிக்கலான கருத்துக்கள்மூலம் அல்லது தவறாகப் புரிந்து கொண்ட கேள்விகள்மூலம் நிகழும் மாதிரி முறை சாராப் பிழைகள் (non-sampling errors) குறைக்கப்படும் விதமாக, அங்கங்கே விளக்கக் குறிப்புகள் (Explanatory notes), தகுந்த உதாரணங்கள் (Specific Examples) ஆகியவற்றைத் தந்து உதவ வேண்டும். பலவகைச் சொற்களும், வறையறைகளும் ஆய்வுக்கு எந்த அளவில் உதவுகின்றன என்று விளக்கமாகக் கூற வேண்டும். ஆய்வாளருக்கு எழும் பல ஐயப்பாடுகள் முன் கூட்டியே தெளிவுபடுத்திக் கொள்ளப்படவேண்டும். அப்போதுதான் ஆய்வாளரிடை ஒற்றுமை இருக்கும். அஞ்சல் வழியாகப் பட்டியல்கள் அனுப்பப்பெறின், சுருக்கமாகவும் புரியும்படியும் விளக்கங்களும் குறிப்புக்களும் தரவேண்டியது அவசியம். நேரிடைத் தொடர்பிலும், சிக்கலற்ற, தெளிவான விளக்கங்கள் கொடுக்கப்படல் வேண்டும்.

விவரங்களைப்பெற உதவும்பொருட்டு, வினாக்களைத் தொகுக்கையில் சிலவற்றைக் கருத்தில் கொள்ளல் வேண்டும். அவற்றைக் கீழ்க்கண்டபடி குறிப்பிடலாம் :

1. எளிமையாகவும் சுருக்கமாகவும் விடையளிக்கத்தக்கவாறு வினாக்களை அமைக்கவேண்டும். குறைந்த அளவு கல்வி அறிவு பெற்றவர்களாலும் அவை புரிந்துகொள்ளப்படின் முழு அளவில் பதில்களைப் பெறலாம். பெரிய பெரிய வாக்கியங்களாக இருப்பின் பதிலளிப்போர் பொறுமை குறைய, ஆய்வின் அடிப்படை நோக்கமே குலைய ஏதுவாகிறது.

2. 'ஆம்' அல்லது 'இல்லை' என்று அல்லது 'தெரியாது' ("do not know") என்று திட்டவட்டமாகப் பதில் தரும்படியாக வினாக்களை எழுப்பவேண்டும். உதாரணம் : "பணி புரிபவரா? ஆம்—இல்லை"

"சென்ற தேர்தலில் வாக்களித்தவரா ?—ஆம்—இல்லை"

3. ஒரு குறிப்பிட்ட பகுதியில்மட்டும் அன்றாட வழக்கில் உள்ள சொற்களையும், சொற்றொடர்களையும் (Expressions) தவிர்த்தல் அவசியம். கலைச் சொற்களை (Technical Words), நன்கு அறிமுகமான மக்களிடமே பயன்படுத்துவது நன்று.

4. எண்ணிக்கையில் குறைந்த அளவு (Minimum) வினாக்கள் அமைவது சாலச் சிறந்தது. ஆய்வு நோக்கத்துக்கு சற்றும் பொருந்தாத வினாக்களை நீக்கவேண்டும். குறிப்பாக, அஞ்சல் வழி ஆய்வில், வினாக்களின் எண்ணிக்கை கூடுமான வரையில் குறைவாக இருத்தல் அவசியம்.

5. நுண்ணறிவின் அடிப்படையில், தர்க்க ரீதியில் (Logically) வினாக்களைத் தொகுப்பது உடனடியான பதில்களைப் பெற உதவுகிறது. முந்திய கேள்விக்கு உள்ள பதிலின் இயல்புக்கு ஏற்பு அடுத்து வரும் கேள்வி அமைந்திருக்கவேண்டும். பொதுவான வினாக்களை (பெயர், இனம், வயது, தொழில் போன்றவை) முதலில் அமைப்பதானது சிறந்த தொடக்கமாகும். கால அடிப்படையிலும் வினாக்கள் தொடர்புபெற்றிருக்க வேண்டும் (Time Sequence). ஒரு குறிப்பிட்ட கருத்தைப்பற்றிய வினாக்கள் அனைத்தும் ஒரு பகுதியில் அமைவது பதில்களை எளிதாகப் பெற வாய்ப்பளிக்கும்.

6. பதிலளிப்போரது சொந்த விவரங்கள்பற்றியோ, அவருக்கு எரிச்சல் அளிக்கத்தக்கதாகவோ, பதில் தர இயலாத வண்ணமோ அமையும் வினாக்கள் கூடியவரையில் தவிர்க்கப்பட வேண்டும். உதாரணமாக, பெண்களிடம் கேட்கப்படும் வயது விவரம், நபர்களின் சொத்து விவரம் இவற்றைப் பற்றி, மறைமுகமான வழிகளில் விவரங்கள் சேகரிக்கப்படல் வேண்டும்.

7. பயன்படுத்தும் சொற்களை மற்றும் அளவு முறைகளைத் தெளிவாகக் குறிப்பிடவேண்டும். "சமீபத்திய பிறந்த தினப்படி வயது என்ன ?" என்பது போன்ற வினாக்கள் இவ் வகையைச் சாரும். தேவையான விளக்கங்களை ஆய்வாளர் தர முன்வரவேண்டும்.

8. ஆய்வாளரின் மதிநுட்பத்தால், சில விவரங்களைக் கூடிய வரையில் சரியாகப் (correctly) பெறமுடிகிறது. துணைக் கேள்விகளை எழுப்பி, சில விவரங்கள் அறியப்படலாம். ஒரு நபரின் வருமானம் பற்றி அறிய,

"செய்யும் தொழில் என்ன ?"

"அதில் ஈட்டும் வருமானம் என்ன?"

"நிலம், வீடு போன்ற சொத்துக்கள் உண்டா ?"

"விளை நிலமாயிருந்தால் அதில் கிட்டும் வருமானம்"

"வேறு வழிகள்"

போன்ற வகையில் பல கேள்விகளைக் கேட்டு, வருமானத்தை மதிப்பிட்டு அறியலாம்.

9. ஒரு கேள்விக்குப் பல பதில்களை வினாத் தொகுதியிலேயே தருவதன்மூலம், ஏதேனும் ஒன்றைச் சரியாகத் தெரிவு செய்ய பதிலளிப்போருக்கு வாய்ப்பு கிட்டுகிறது. உதாரணம். போக்குவரத்து (transport) பற்றிய ஆய்வில் ஒரு கேள்வி :

"தாங்கள் அலுவலகம் செல்வது எப்படி ?

☐ சொந்த வாகனத்தில்

☐ பேருந்தின் மூலம்

☐ மிதிவண்டி மூலம்

☐ கால் நடையாக

☐ வேறு வழியில்.

10. நடைமுறையில் ஏற்படக்கூடிய இடர்ப்பாடுகளையும் சிக்கல்களையும் முன்கூட்டியே திட்டமிட்டு அறிந்து அதற்கேற்ப செயல்படுவது மிகவும் சிறந்தது. பல பிரச்சினைகளை இதன் மூலம் தீர்க்க வழி ஏற்படுகிறது.

எனவே, ஆய்வின் மேல் வரும் கட்டங்கள் அனைத்துமே வினாப்பட்டியலின் திட்ட அமைப்பினைப் (designing) பொறுத்தவை என்று கூறுவது எவ்வகையிலும் மிகையாகாது. ஆய்வின் வெற்றியே இதில்தான் அடங்குகிறது எனலாம்

வினாப்பட்டியல் 1

ஒரு உதாரணம்

கல்லூரி மாணவரது செலவினங்களின் (expenditure) தன்மையை அறிய அமைக்கப்படும் வினாத் தொகுப்புப் பட்டியல் (questionnaire) :

(இரகசியமானது)

(Confidential)

1. **பொதுவான விவரங்கள் :**

(அ) பெயர் : ...

(ஆ) வகுப்பு : ...

(இ) பிரிவு : ...

(ஈ) வயது : ... பிறந்த தேதி : ...

(உ) தந்தையின் பெயர் : ....

(ஊ) செய்யும் தொழில் : ....

(எ) ஆண்டு வருமானம் : ரூ ......

(ஏ) வேறு வழிகளில் வருமானம்
 (i) நிலங்களில் ரூ : ...
 (ii) சொத்துகள்மூலம் ரூ : ...
 (iii) இதரவகைகள் ரூ : ...

(ஐ) கூடப்பிறந்தவர்கள் : ... சகோதரர்கள் : ...
(எண்ணிக்கைகளைக் குறிக்க) சகோதரிகள் : ...

2. **கல்வி பயிலும் நிலை :**

(அ) சம்பளச் சலுகை பெறுபவரா ? ...
'ஆம்' எனில், எத்தகைய சலுகை ? முழு அளவில் / அரைப் பகுதி / இதர வகை.

(ஆ) உபகாரச் சம்பளம் பெறுபவரா ? ஆம் / இல்லை.
(Scholarship)
'ஆம்' எனில், பின்வரும் விவரங்களைத் தருக :
(i) எவ்வாறு உதவி பெறப்படுகிறது ? தமிழக அரசு / மத்திய அரசு / கடனுகை / வேறுவிதத்தில்
(ii) மொத்தம் பெறும் உதவித்தொகை : ரூ ...

(இ) ஏதேனும் இதர உதவிகள் பெறுவதாயிருந்தால் அவற்றைக் குறிப்பிடுக. எவ்வகையில் பெறும் தொகை ரூ

(ஈ) வீட்டிலிருந்து மாதந்தோறும் பெறும் தொகை : ரூ ...

(உ) மொத்தம் பெறும் பணத்தொகை : ரூ ...

3. **செலவு இனங்கள் பற்றிய விவரம் :**

செலவு வகை (சமீப மாதத்தில்)	அளவு	செலவான தொகை ரூ.	பை
(i) உணவு விடுதிக்கு			
(ii) ஓட்டல்களுக்கு			
(iii) அருந்திய பான அளவு			
காபி ... (கோப்பைகளில்)			
டீ ...			
இதர பானங்கள்			
(iv) சலவைக்கு			
(v) மருத்துவச் செலவு			
(vi) மருந்துகளுக்கு ஆன செலவு			
(vii) சோப்பு			
(viii) ஹேர் ஆயில்			
(ix) பேஸ்ட்			
(x) பவுடர்			
(xi) முடி வெட்டுதல்			
(xii) சிகரெட்டுகள்			
(xiii) வெற்றிலை, பாக்கு, புகையிலை			
(xiv) பிளேடுகள்			
(xv) தினசரிப் பத்திரிகைகள்			
(xvi) வாரப் ,,			
(xvii) மாதப் ,,			
(xviii) இதரப் ,,			
(xix) பாடப் புத்தகங்கள்			
(xx) நாவல்கள்			
ஆங்கிலம்			
தமிழ்			
இதர வகை			

(xxi) நோட்டு புத்தகங்கள்

(xxii) இங்க்

(xxiii) பஸ் செலவு

(xxiv) ரயில் வண்டிச் செலவு

(xxv) தர்மங்கள்

(xxvi) பொது நிறுவனங்கட்குக் கொடுத்தது.

	அளவு	மாதச்செலவு
(xxvii) கடனுகக் கொடுத்த பணம்		ரூ. பை
(xxviii) சினிமா செலவு		
(xxix) விருந்தினர் வகையில் செலவு		
(xxx) இதர வகைச் செலவு		
மொத்தம்		

4. சென்ற ஆண்டு, பின்வரும் வகைகளுக்கு ஆன மொத்தச் செலவு விவரம் :

(i) பாடப் புத்தகங்கட்கு ரூ.

(ii) இதரப் புத்தகங்கட்கு ரூ.

(iii) அடிப்படைத் தேவைகட்கு ரூ.

(iv) எதிர்பாராத செலவு ரூ.

(v) தொலைந்துபோன பணம் ரூ.

5. மேற்கண்டவற்றில் இடம் பெறுத இதர விவரங்கள் (ஏதேனும் இருப்பின்)

வினாப்பட்டியல் 2

சமூகப் பொருளாதார ஆய்வு பற்றிய வினாத்தொகுதி (Schedule for Socio-Economic Survey)

வசிக்கும் நகர் : தொகுதி :

(Ward) :

தெரு :

வீட்டு எண் :

I பொதுவான விவரங்கள் :

1. குடும்பத் தலைவர் பெயர் : ஆண்/பெண்

2. வயது : பிறந்த தேதி :

3. செய்யும் தொழில் :

4. வருமானம் (ஒரு ஆண்டுக்கு) :

5. வருமானவரி கட்டுபவரா ?

'ஆம்' எனில், ஒரு ஆண்டுக்கு கட்டும் வருமானவரி அளவு } : ரூ

II குடும்பம், மற்றும் வசிக்கும் சூழ்நிலை பற்றி :

1. குடும்பத்தில் வசிக்கும் நபர்கள் எண்ணிக்கை :

மொத்தம் ...

ஆண்கள் (18 வயதுக்கு மேல்)

பெண்கள் (18 வயதுக்கு மேல்)

குழந்தைகள் (18 வயதுக்கு கீழ்) ஆண் : பெண் :

2. எழுதப் படிக்கத் தெரிந்தோர்: ஆண் : பெண் :

3. பள்ளி வரை கல்வி பயின்றோர் : ஆண் : பெண் :
(School level)

4. பட்டப்படிப்பு முடித்தோர் : ஆண் : பெண் :

5. சம்பாதிக்கும் இதர நபர்கள் : ஆண் : பெண் :

6. வசிக்கும் வீட்டின் தன்மை : (பொருத்தமானவைகளை 'V' என்று குறிப்பிடவும்)

(i) சொந்த வீடு

(ii) வாடகை வீடு

(iii) ஓட்டு வீடு

(iv) 'டெரஸ்' வீடு (terraced)

(v) மின்வசதியுடன் கூடியது

(vi) இதர வசதிகள் (குறிப்பிடவும்).

7. வசிக்கும் பகுதியில் பின்வரும் வசதிகள் உண்டா :

(i) பாதாள சாக்கடை (drainage)

(ii) அவ்வப்போது சுத்திகரித்தல் :

(iii) தெரு விளக்குகள் :

(iv) தெருத் துப்புரவு :

(v) இதர வசதிகள் :
(குறிப்பிடுக)

2. வீடு அமைந்திருக்கும் இடம் :

(i) நெடுஞ்சாலை :

(ii) சந்து :

(iii) இரண்டும் இல்லை :
(குறிப்பிடுக)

III. வரவு—செலவு விவரம் :

1. (அ) வருமானம் : மொத்தம்
(ஒரு ஆண்டுக்கு) ரூ :

(i) வீடு அல்லது நிலம் மூலம் ரூ :

(ii) இதர சொத்துகள் மூலம் ரூ :

(iii) தொழில் மூலம் ரூ :

(iv) இதர வகையில் ரூ :

(ஆ) குடும்பச் செலவு, சராசரியாக ஒரு மாதத்திற்கு (சென்ற ஆண்டின் அடிப்படையில்)

(i) உணவுப் பொருட்கள் ரூ :

(ii) வீட்டு வாடகை அல்லது பழுது பார்க்கச் செலவு ரூ :

(iii) எரிபொருள் மற்றும் ஒளிவசதி ரூ :

(iv) துணி வகைகள் மற்றும் வெளுத்தல் இவற்றிற்கு ஆகும் செலவு ரூ :

(v) மருந்து வகைகள், டாக்டர் போன்றவை ரூ :

(vi) முடி வெட்டுதல் போன்றவை கட்டு ரூ :

(vii) மதச் சடங்குகள் ரூ :

(viii) கேளிக்கை வகைச் செலவு ரூ :

(ix) இதர செலவு வகை ரூ :

மொத்தம்: ரூ :

2. குடும்பச் செலவு போக சேமிப்பு விவரம் ரூ :

3. சேமிப்பைப் பயன்படுத்தும் விதம் ரூ:

4. பற்றுக் குறையினைச் சமாளித்த விதம் ரூ :

5 கடன் வாங்குவாரா ?

'ஆம்' எனில் எவ்வாறு ? அடிக்கடி
அவ்வப்போது
மாதக்கடைசியில்
தெரியவில்லை.

6. பொழுது போக்குகளில் விருப்பம் உண்டா?

7. இதர விவரங்கள் இருப்பின், குறிக்க :

5. அளவெடுப்பு மேற்கோள் மற்றும் அறிவிக்கும் காலவரைகள் (Survey, Reference and Reporting Periods)

விரிவாக உரைத்திடின், இங்கு

(i) ஆய்வின் கால வரையறை

(ii) குறிப்பீட்டுக் கால அளவு

(iii) அறிக்கை சமர்ப்பிக்கும் காலவரை

ஆகியவற்றை நன்கு தெளிவாகத் தீர்மானிப்பதில் தனியான கவனம் செலுத்தப்படவேண்டும்

மேற்கோள் கால அளவானது (reference period), சேகரிக்கப்படும் அனைத்துக் கூறுகளின் விவரமும் குறிப்பிடப்பெறும் கால அளவைக் காட்டும். ஆய்வின் அடிப்படை நோக்கத்தைப் பொறுத்தே இது நிர்ணயிக்கப்படும், அறிவிக்கும் காலவரையானது (reporting period), கண்டறியப் பெறும் விவரக் கூறுகளையும், ஆய்வு நடைபெறும் சூழ்நிலையையும்(condition) பொறுத்து அமையும். பலவகைக் கூறுகளுக்குத் (items) தக்கபடி, இவ்விரண்டு கால வரையறைகளும் வெவ்வேறாக அமைதல்வேண்டும் மாதிரி முறை சாராத (non-sampling) பிழைகளை நன்கு பரிசீலனை செய்து அறிந்த பின்னரே, அறிவித்தற்கான கால அளவை நிர்ணயிக்கவேண்டும். வேலை வாய்ப்பின்மை போன்ற ஆய்வுகளில், ஒரு வார காலமானது அறிவிக்கும் காலமாக ஏற்றுக்கொள்ளப்படுகிறது. இது நிலையானதாக இருத்தல் அவசியமன்று.

பருவகால மாறுபாடுகட்கு (seasonal fluctuations) உட்பட்ட விவரக் கூறுகளைக் கருதுகையில் ஆய்வு நடத்தும் காலமும், குறிப்பீட்டுக் காலமும் (survey & reference periods) ஒரு முழு ஆண்டாகத் தேர்ந்தெடுத்து, பிரதி மாதமும் (அல்லது ஒவ்வொரு பருவத்திலும்) விவரங்களைச் சேகரித்துப் பெறவேண்டும். ஒவ்வொரு அலகிற்கும் (unit) அறிவித்தற் காலமானது (reporting period), நிலையாகவோ (same) அல்லது மாறுபட்டோ இருக்கலாம். பருவகால மாறுபாட்டை அறியவேண்டின், ஒரே தொகுதியில் அமைந்த மாதிரிக் கூறுகளைக் (sample units) கருதலாம். மாறாக ஒரு ஆண்டில் உள்ள சராசரி நிலையை அறிய, வெவ்வேறு தொகுதிக் கூறுகளை ஆய்தல் வேண்டும்.

ஆனால் பெரிய அளவில் நடத்தப்பெறும் ஆய்வினில், கால அளவினை நீட்டிக்கவேண்டிய நிலை ஏற்படும். இத்தகைய நிலைகளில், நகரும் அறிவித்தற் காலத்தினை (moving reporting period) பயன்படுத்தவேண்டும். உதாரணமாக, ஓர் ஆண்டுக் கால ஆய்வில், சனவரி முதல் டிசம்பர் வரை கருதப்பட்டால், நிகழ்காலத்தில் ஒரு ஆண்டுக்காலக் குறிப்பீட்டினை (reference) அளிப்பதற்கேற்ப, போன ஆண்டு டிசம்பர் முதல் இந்த ஆண்டு நவம்பர் வரையிலான ஒரு ஆண்டு எடுத்துக்கொள்ளப்படும்.

குறிப்பிடும் காலமானது (reference period) அனைத்துக் கூறுகட்கும் ஒன்றானதாக அமைந்திடும்பட்சத்தில், சில விவரங்களைச் சரிபார்ப்பதற்கும், வெவ்வேறு உறுப்புகளிடையே (items) உள்ள தொடர்பை அறிந்திடவும் ஏதுவாகிறது. நீட்டிக்கப்படும் (staggering) ஆய்வுகளில், ஆய்வுக்கால அளவை பல சிறு பகுதிகளாகப் பிரித்து, (ஒரு மாதம், இரு மாதம்...என்று) ஒவ்வொரு பகுதிக்கும் ஏற்ற உட்கூறு மாதிரியைக் (subsample) கருதி, ஒவ்வொரு காலத்திற்கும் ஏற்ற மதிப்பீடுகளை அடையலாம்.

6. மாதிரிச் 'சட்ட' வரையறை (Sampling Frame Determination)

மாதிரி ஆய்வுமுறைச் சிறப்புக்கு, வரையறுக்கப்படும் மாதிரிக் கூறுகள் (sample units) அடங்கிய மாதிரி 'சட்ட' அல்லது வரம்பு அமைப்பும் ஒரு காரணமாக விளங்கும். இது நடைமுறைக்கு ஏற்றவாறும், துணை இயல்புகள் (auxiliary characteristics) பற்றிப் போதுமான விவரங்கள் அளிக்கும் தன்மையானதாகவும் இருத்தல் இன்றியமையாதது. இத்தகைய மாதிரிச்சட்டங்கள் பல நிலைகளில் கிடைக்க இயலாததால் பிரச்சினைகள் ஏற்படு

கின்றன. இவை கிடைக்கும் இடங்களில், விவரங்கள் விடப் படாமலும், ஒரே விவரம் திரும்ப வராமலும், மற்றக் குறைகள் ஏற்படாதும் தடுக்கும் வகையில் சில நடவடிக்கைகளை எடுக்க வேண்டி அமையும்.

விரிவானதொரு விவரப்பட்டியல் கொடுக்கப்படும் நிலையில், எந்த ஆய்வாளரும். பகுதிகளாகப் பிரிப்பதற்கும் (stratification), ஒதுக்கீடு செய்தற்கும் (allocation), தேர்ந்தெடுப்பதற்கும் (selection), மதிப்பீடு செய்வதற்கும் (estimation) பெருமளவில் உதவக்கூடிய பல்வேறு வகையான துணை விவரங்களைப் பெறும் சாத்தியக் கூறினை ஆராயவேண்டும். மாதிரித் திட்ட அமைப்பிற்கு (sampling design) மிக உதவும் இத்தகைய துணை விவரங்கள் (auxiliary information) முற்றிலும் சரியானவையாக இருத்தல் அவசியமன்று. தேவையான துணை விவரங்களுடன் அமைக்கப் படும் சிறந்த சட்ட வரையறை (master frame) ஆனது, ஆய்வில் அதிகப்படி பலனை அளிக்கத்தக்கது.

பலகட்ட மாதிரி முறையில் (multistage sampling), ஒவ்வொரு கட்டத்திலும் மேற்குறிப்பிடப்பட்ட சிக்கலானது எழுகிறது. உயர்மட்ட (higher stage) உறுப்புகளினைச் சார்ந்த மாதிரிச் 'சட்ட' (frame) அமைப்பானது மிகவும் நிலையான (stable) தன்மையுடைத்ததாகக் காணலாம். அதாவது, நகரங்கள், கிராமப் புறங்கள் என்ற வகையில் செய்யப்படும் பாகுபாடு, கீழ்மட்டத்தில் நிறுவனங்கள், குடும்பங்கள் என்ற வகைப் பாகுபாட்டைக் காட்டிலும், நிலையான 'சட்ட அமைப்புகளைக் கொண்டிருப்பதைக் காணலாம். ஆட்சிமுறைக் குறிப்பேடுகளிலிருந்தும் (administrative records), முழுக் கணிப்பின் முடிவுகளிலிருந்தும் (census) இத்தகைய சட்ட அமைப்புகளை உருவாக்குதல் எளிது. ஆனால், இறுதிக்கட்ட (final stage) மாதிரிக் கூறுகளுக்கான 'சட்ட' வரையறையினை அமைத்தல் சிரமமான காரியமே. முதலிலிருந்து தொடங்கிப் பட்டியல் வகுக்கவேண்டும். இதோடு இணைந்த துணை விவரங்களும் பெறப்பட வேண்டும். இதன்மூலம், இறுதியில் ஏற்படும் வேறுபாடுகளைக் குறைத்து, திட்பமான மதிப்பீடுகளைப் பெறுவது சாத்தியமாகிறது. அமைக்கப்படும் 'சட்ட' அல்லது 'வரம்பு' வரையறையினைப் பொறுத்து, கையாளவேண்டிய மாதிரியின் திட்ட அமைப்பு (design) உருவாவதை அடுத்த பகுதியில் காணலாம். மாதிரி முறை எனும் சித்திரம் அழகாக வடிவம் பெற, மாதிரி 'சட்ட' வரையறை எனும் சுவரானது நன்கு அமையவேண்டும் என்று குறிப்பிடலாம். முதல் கட்டத்தில் ஏற்படும் தவறு அல்லது பிழை (error) இறுதிவரையில் அதிகரிக்கப்பட்டு, இறுதி முடிவுகள் பாதிக்கப்படுகின்றன.

7. (அ) மாதிரி முறையின் திட்ட அமைப்பு (Sampling Design)

ஆய்வுச் செலவு, செயல்முறை, சூழ்நிலை போன்றவை றுக்கு ஏற்றவகையில் ஒரு சிறந்த மாதிரி முறை வடிவினை (optimum sampling design) வரையறை செய்வது சாத்தியமல்ல இருக்கின்ற பல்வேறு வழி முறைகளில் சிறந்த ஒன்றினை தெரிவு செய்வதே நடைமுறையில் இயலும். பொதுவாக, நிர்ணயம் செய்யப்பட்ட அனுமதிக்கக்கூடிய பிழை எல்லையுடன் (permissible error), ஆய்வின் மொத்தச் செலவினைக் குறைக்கும் விதமாக, அல்லது, கொடுக்கப்பட்ட ஆய்வுச்செலவுக்கேற்ப மதிப்பீடுகளின் பிழை எல்லைகளின் அளவைக் குறைத்திடும் விதத்தில், தக்க தொரு மாதிரி வடிவமைப்பைத் தேர்வு செய்கிறோம். இரு முறை களிலும், மாதிரிமுறை சார்பற்ற பிழைகள் (non-sampling errors) கட்டுப்படுத்தப்பட்டு இருக்கவேண்டும். நடைமுறைச் சூழலுக்கு ஏற்றதாக, தெரிவுசெய்யப்பட்ட மாதிரி முறையானது அமைவது இன்றியமையாத ஒன்றாகும்.

பெரிய அளவில் நடத்தப்பெறுகின்ற மாதிரி அளவெடுப்பு களில், (sample surveys) ஒரு பகுதி—பிரிக்கப்பட்ட (படுகை பற்றிய), ஒருகட்ட அல்லது பலகட்ட (stratified unistage or multistage) மாதிரி முறைத் திட்ட அமைப்புதான் பின்பற்றப்படு கிறது. கருதப்பட்ட இயல்புகளுடன் தொடர்புடைய பல துணை விவரங்கள் கிடைப்பதைப் பொறுத்து, படுகை முறை (stratification) விரிவாகப் பயன்படுத்தப்படும். முதலில், உத்தம அளவினதாக (optimum) சில பிரிவுகளை அல்லது படுகைகளை (strata), ஒரு படுகை மாறியினை (stratificative variable) பயன் படுத்தி அமைக்கிறோம். மற்றொரு படுகை மாறியின்மூலம், அமைத்த படுகைகளின் உட்பிரிவுப் படுகைகளை (sub-strata) உருவாக்குகிறோம். ஒவ்வொரு படுகை அல்லது பிரிவு அமைப் பும் ஒரு ஆய்வாளரது 'ஆய்வு மையம்' (investigation zone) என்று குறிப்பிடப்படும். இவ்வாறு அமைக்கப்பட்ட பிரிவு அல்லது படுகைகளுக்கு ஏற்ப ஒதுக்கீடு செய்யப்படும்(allocation) மாதிரி அளவுகள் (sample size), துணை மாறிகள் (auxiliary variables) பற்றிய விவரங்களைப் பொறுத்து அவற்றின் அடிப் படையில் அமைகின்றன.

பின்பு, அமைக்கப்பட்ட பிரிவுகள் அல்லது படுகைகளி னின்று, அலகுகளைத் தெரிவு செய்ய, முறையுடை அல்லது ஒழுங்கு மாதிரி முறை (systematic sampling) கையாளப்படு கிறது. மாறுபடும் வாய்ப்பு (varying probability) முறைபற்றிய மாதிரித் தேர்வு முறைகளில், அளவினை (size) அறிவது முக்கிய

கட்டமாகும். எனினும், கருதப்பட்ட அளவுகள் ஏறத்தாழ சம மாக அமைந்திருப்பின் (மக்கட் தொகைக் கணிப்பு ஆய்வில், கருதப்படும் ஆய்வுத் தொகுதிகள் (blocks) போல, சமவாய்ப்பு அடிப்படையில் (equal probability basis) ஒழுங்கு முறை மாதிரிகள் தெரிவு செய்யப்படலாம். மதிப்பீட்டுக் கட்டத்தில், விகித அல்லது உடன்போக்கு (ratio or regression) மதிப் பீடுகள்மூலம், மற்ற அளவு (size) மதிப்புகளும் கருதப்பட லாம்.

இரு—கட்ட மாதிரித் தேர்வில் (two-stage sampling), முதல் கட்டத்தில் தேர்ந்தெடுக்கப்பட்ட கூறுகளிலிருந்து, இரண்டாம் கட்டக் கூறுகளை ஒழுங்குமாதிரி முறைப்படி (systematic), 'சட்ட' அமைப்பை முழுவதுமாகவோ அல்லது அதிலிருந்து பெறப்பட்ட 'துணை சட்ட' அமைப்பையோ பயன்படுத்தித் தெரிவு செய்ய லாம். இதே முறையினைத் தொடர்ந்து பல கட்டங்களிலும் பயன் படுத்தலாம். மேலும், பல கட்டங்களிலும், திரட் கூறுகளைப் (cluster) பயன்படுத்துவதானது செலவைக் கட்டுப்படுத்தும் என்று குறிப்பிடுதலும் அவசியமாகிறது.

ஒன்றுக்குள் ஒன்று ஊடுருவும் வண்ணம் (interpenetrating) துணை—மாதிரிகளாய் (sub—sample) மாதிரித் தேர்வு அமையப் பெறின், முழுமைத் தொகுதி அளவைக்கு (population parameter) ஏற்ற ஏற்கத்தக்க மதிப்பீடுகளை ஒவ்வொரு மாதிரியும் அளிக்க வல்லதாகும். இம்முறையில், மாத்ரி முறை சார்ந்த பிழைகளை (sampling errors) மதிப்பிட்டு அறிவது எளிதாகிறது. முதல் கட்டத்தில், வேறுபடும் வாய்ப்பு நியதிப்படி, ஒழுங்கு முறையில் மாதிரிகள் எடுக்கப்படும் திட்டமுறைகளில் இது முக்கியமாகிறது. பின்வரும் இருவகையன முறைகளில் மாதிரித் தேர்வு அமையக் காணலாம்.

7. (ஆ) மாதிரித் தேர்வின் வழிமுறைகள் (Sampling Techniques)

ஒரு மாதிரியினைத் தேர்ந்தெடுப்பது பல வழிமுறைகளைப் (method) பின்பற்றியதாக அமையலாம். பொதுவாக அவற்றை இரு கூறுகப் பிரித்து ஆராயலாம்.

(1) நிகழ்தகவின் அல்லது வாய்ப்பின் அடிப்படையில் அமைந்த மாதிரி முறைகள் (Probability Sampling).

(2) நிகழ்தகவின் அடிப்படையில் அமையாத மாதிரி முறைகள். (Non probability Sampling).

முதலாம் வகையில், நிகழ்தகவின் அடிப்படையில் மாதிரிகளைத் தெரிவு செய்வது பல்வேறு வழிகளில் செயல்படும். இம்முறையினைக் கண்டறிந்து நடைமுறையில் பயன்படுத்தியவர் டெமிங் (Deming) ஆவார் (1940). இதில் அமைந்த பல்வேறு கூறுகளாவன :

(i) சாதாரண ராண்டம் மாதிரி முறை (Random sampling)

(ii) படுகை மாதிரி முறை அல்லது பகுத்துக் கூறெடுக்கும் மாதிரி முறை (Stratitied Sampling)

(iii) ஒழுங்குமாதிரி முறை (Systematic Sampling)

(iv) திரள் மாதிரி முறை (Cluster Sampling)

நிகழ்தகவின் அடிப்படையில் முடிவுகளைத் தீர்மானிக்க இவ் வழிமுறைகள் மிக உதவுகின்றன. நிகழ்தகவுக் கோட்பாட்டைப் பயன்படுத்துவதன்மூலம், மாதிரித் தெரிவில் ஏற்படும் பல இடர்ப்பாடுகள் நீக்கப்படுகின்றன. ஒரு புறச் சார்பற்ற, ஒழுங்காக எடுக்கப்படும் மாதிரிகள், கருதப்பட்ட முழுமைத் தொகுதியின் இயல்பினை முற்றிலும் பிரதிபலிக்கின்றன. மேற்கூறப்பட்ட வழி முறைகளை விரிவாகப் பின்னால் ஆராய முற்படுவோம்.

இரண்டாவது வகையில், நிகழ்தகவு விதிகள் பின்பற்றப்படுவதில்லை. தெரிவு செய்யப்படும் கூறுகளுடன் நிகழ்தகவு (அல்லது வாய்ப்பு) தொடர்புபடுத்தப்படுவதில்லை. எனவே, அடையப்பெறும் முடிவுகளின் நம்பகத் தன்மையின் அளவை நிகழ்தகவு விதிப்படி உறுதி செய்திட இயலுவதில்லை. நிகழ்வு தகவு அடிப்படை மாதிரி முறை கண்டறியப்படுவதற்கு முன்பாக இத்தகைய வழிமுறைதான் பின்பற்றப்பட்டது. இவ் வகையில் அமைந்த சில மாதிரி முறைகள் வருமாறு :

(i) நோக்கத்துடன் எடுக்கப்பட்ட மாதிரி முறை (அல்லது ஒரு தீர்ப்பின் அடிப்படையிலான மாதிரி முறை).

(Judgement Sampling or Purposive Sampling)

(ii) அஞ்சல் வழி வினாப் பட்டியல் முறை (Mailed-Questionnaire)

(iii) ஒதுக்கீட்டு அளவு மாதிரி முறை (Quota-Sampling).

நோக்கத்துடன் எடுக்கப்படும் மாதிரி முறை

இம் மாதிரி முறையானது, நிகழ்தகவின் அடிப்படையில் அமையாமல், ஒரு குறிப்பிட்ட நோக்கத்திற்காக வேண்டுமென்றே தேர்ந்தெடுக்கப்படும் ஒன்றாகும். பிறழ்ச்சிக்கு அதிகமாக வாய்ப்பிருப்பதால், புள்ளிவிவர ஆய்வாளரின் விருப்பு வெறுப்புக்கு ஏற்ப, மாதிரி முறையில் தவறுகள் ஏற்படுகின்றன. ஆய்வு நடத்தும் நபரின் தன்மைக்கேற்ப மாதிரிக் கூறுகள் அமைகின்றன. விவரங்களைத் தவறான வழியில் பயன்படுத்த வழி உண்டு. அரசியல் வாதியின் விருப்பத்திற்கேற்பப் புள்ளி விவரங்கள் கையாளப்படுவது இதற்கு ஓர் எடுத்துக்காட்டு. ஒரு நகரின் வாழ்க்கைத் தரத்தினை உயர்ந்ததாகக் காட்டும் பொருட்டு, வசதியுள்ள மக்களைக் கொண்டு அளவிடப் பெறும் கணிப்பு முறை இந்த வகையினைச் சார்ந்ததாகும். இம் முறையானது, முழுமைத் தொகுதியின் இயல்பினைப் பிரதிபலிக்கும் மாதிரியைத் தராது. பெறப்படும் முடிவுகள் எதிர்பார்த்த அளவு திட்பத்துடன் (precision) அமையா. சொந்த விருப்பு வெறுப்புகளும், ஒரு புறச் சார்புத் தன்மையும் அதிக அளவில் காணப்படும். மேலும், கொடுக்கப்பட்ட முழுமைத் தொகுதியின் இயல்புகள் ஆய்வாளருக்கு நன்கு தெரிந்திருக்கவேண்டும். அப்போது தான் சிறந்த முறையில் விவரங்களைக் கண்டறிய இயலும். அப்படியும், ஒருபுறச் சாய்வினைத் (bias) தவிர்க்க எவராலும் இயலாது. 'தவறு செய்வது மனித இயல்பு' என்பது பொது விதி அன்றோ! தனி ஒரு மனிதனின் தீர்ப்பு அல்லது முடிவுப்படி மாதிரியானது தெரிவு செய்யப்படுவதால், இம் முறை சிறந்ததாக ஆகாது. சிறந்த அனுபவமும் பயிற்சியும் உடைய இருவரது முடிவுகள் எப்போதுமே ஒத்திருப்பதில்லை. இரு ஆசிரியர்கள் தேர்ந்தெடுக்கும் மாணவர்கள், அவரவர் தம் சொந்த விருப்பத்துக்கேற்ப அமைந்திருப்பதைத் தவிர்ப்பது மிக்க கடிமானதொரு செயல். ஒரு குறிப்பிட்ட எண்ணத்துடன் பயன்படுத்தப்படும் இந்தச் செயல்முறையானது, புள்ளியியல் நியதிப்படி எவ்விதத்திலும் சிறந்ததல்ல.

அஞ்சல் வழி வினாப்பட்டியல் முறை (Mailed Questionnaire Method)

இம் மாதிரி முறை, நிகழ்தகவு அடிப்படையில் அமையாது எடுக்கப்படும் முறை. வினாப்பட்டியல் பதிலளிப்போருக்கு அஞ்சல் வழி மூலம் அனுப்பப்பட்டு, பூர்த்தி செய்த பதில்கள் அஞ்சல் மூலம் திரும்பப் பெறப்படுகிறது. பதிலளிப்போர் பலர்

கணிப்பின் (survey) நோக்கத்தை அறியாதபொழுது பதில்களை அளிக்காது விட்டுவிடலாம். இதனால் முழுதும் பூர்த்தி செய்யப்படாத வினாப்பட்டியல்களோ அல்லது பூர்த்தி செய்யப்பட்ட வினாப்பட்டியல்களோ ஆய்வாளருக்குத் திரும்ப அனுப்பப்படாது போகலாம். இத்தகைய பதிலின்மை (non- response) சாதாரணமாக, 20% லிருந்து 25% வரை நேரிடுகிறது. இவ்வாறு பதிலின்மை நேரிடும் தறுவாயில், அஞ்சல் வழி அல்லது தொலைபேசி அல்லது நேரடித் தொடர்பு கொண்டாவது தேவையான விவரங்களைப் பதிலளிப்போரிடமிருந்து பெற முயற்சி செய்ய வேண்டும். அவ்வாறு முயற்சி செய்தும் பதிலளிப்போர் சிலர் விவரங்கள் தர மறுத்துவிட்டால், அவர்களை நீக்கிவிட்டு, பதிலளித்தவர்களை மட்டுமே கணக்கில் எடுத்துக்கொள்ள வேண்டும். இத்தகைய அஞ்சல் வழி வினாப்பட்டியல் முறையில் பதிலளிப்போர் தொகுப்பு, பொதுவாக எழுதப்படிக்கத் தெரிந்தவர்களாக (literates) இருப்பின் சிறந்த பலனைத் தரும். இத்தகைய முறையில் வினாக்களின் குறைவான எண்ணிக்கை, வினாக்களை எழுப்பும் முறை, சாதாரண மனிதன் எளிமையில் புரிந்து கொள்ளும் முறையில் அமைவது வினாக்களுக்கான பல்வேறு தக்க பதில்கள் தருவது ஆகியவை அவசியம். இம் முறையானது பெரும்பாலும் பின்பற்றப்படுவதில்லை.

ஒதுக்கீட்டு அளவு மாதிரி முறை (Quota Sampling)

நிகழ்தகவு அடிப்படையற்ற மற்றொரு முறை, ஒதுக்கீட்டு முறையாகும். இதன்படி, ஒரு தோராயமான (crude) முறையினைப் பின்பற்றி, மக்கட் தொகையினை வயது, இனம், வருமானம், மதம் ஆகியவற்றின் அடிப்படையில் பிரித்து, ஒவ்வொரு ஆய்வாளருக்கும் ஒரு மக்கட் பிரிவினை ஒதுக்கீடு செய்து இத்தகைய பிரிவுகளிலிருந்து தேர்ந்தெடுக்கும் மாதிரிகள் பயன்படுத்தப்படும். நபர்களைத் தேர்ந்தெடுத்து மாதிரியினை நிர்ணயிப்பது ஆய்வாளரின் விருப்பத்தை முற்றிலும் பொறுத்தது. மொத்த மக்கட் தொகையானது, உதாரணமாக, ஆண்—பெண் என இரு வகையாகப் பிரிக்கப்பட்டு ஒவ்வொரு பகுதியிலும் ஒரு மாதிரியினைத் தெரிவு செய்யலாம். இத்தகைய பண்புகள் அல்லது இயல்புகள் (characteristics) ஒதுக்கீட்டுக் கட்டுப்பாட்டுக் காரணிகள் (quota control factors) எனப்படும். 'மக்கள் கருத்தறியும் ஆய்வு' (public opinion poll) கணிப்புகளில் இம் முறையானது அதிக அளவில் பயனாகிறது. இதன்மூலம், எளிதான முறையில் குறைந்த செலவில், வெகு விரைவில் மாதிரியானது பெறப்படுகிறது. இது ஒரு குறிப்பிடத்தக்க அனுகூலமாகும்.

ஆனால், கிடைக்கும் மாதிரிக் கூறானது முழுமைக் கூற்றின் இயல்பினை முற்றிலும் பிரதிபலிக்காது. இதன்மூலம் பெற்ற மதிப்பீட்டு அளவைகளின் நம்பகத்தன்மையினையும், திட்ப அளவினையும் (reliability and precision) நிகழ்தகவு விதி அடிப்படையில் நிர்ணயிக்க இயலாது. கிடைக்கும் மாதிரியானது எந்த வகையிலும் கட்டுப்படுத்தப்படாததால், ஒருபுறச் சாய்வின் அளவு அதிகமாக இருக்கும். மேலும் முழுமைக் கூற்று அளவையின் 'நம்பிக்கை இடைவெளி'யினை (confidence interval) வரை செய்ய இயலாது. எனவே, புள்ளியியல் ஆய்வு முறைகளில் இத்தகைய, நிகழ்தகவு சாராத மாதிரி முறைகள் ஏற்கத்தக்க முடிவுகளை அளிக்காததால், அதிக அளவில் பயன்படுத்தப்படுவதில்லை. இத்தகைய வழிமுறைகளில், மனித இயல்புபற்றிய காரண விளைவு அதிகரித்துக் காணப்படும். எனவே ஒவ்வொரு நிலையிலும் ஏற்படும் தவறுகளின் எண்ணிக்கை, ஒருபுறச் சாய்வின் விளைவால் அமையும். இவற்றைக் கண்டறிந்து முற்றிலும் நீக்கிவிடலாம்; எனினும் அது எளிதான செயல் அல்ல. சில இடங்களில் மட்டும் இம் முறைகள் பின்பற்றப்படுகின்றன.

7. (இ) நிகழ்தகவு சார்ந்த மாதிரி முறைகளின் பயன்கள் (Advantages of Probability Sampling)

மாதிரிக் கணிப்பு முறைகளில் நிகழ்தகவுக் கோட்பாடுகள் பெரும் அளவில் உதவுவதைப் பின்வருமாறு எடுத்துக் கூறலாம்:

(i) அறிவியல் அடிப்படையில் மாதிரி வழி முறைகள் (sampling methods) அமைவதால், அவற்றில் நிகழ்கின்ற பிறழ்ச்சியின் (bias) தன்மையினை அளவிடவும் அதைப்பற்றி மேலும் பலவிதங்களை அறிந்திடவும் இயலும்.

(ii) அடுத்து, மாதிரித் தேர்தற் பிழைகளை (sampling errors) அளவிட்டு அறிந்திடவும், இப் பிழைகளில் நிகழ்வுத் தன்மையினை கண்டறிந்திடவும் நிகழ்தகவு விதி மிகவும் பயன்படுகிறது.

(iii) மேலும் பல்வகைப்பட்ட மாதிரிக் கணிப்பு முறைகள் மூலமாகப் பெற்ற மதிப்பீடுகளின் தரப்பிழையினை (standard error) அறிய இயலுவதால், ஒவ்வொரு வழி முறையினது நோக்கத்தையும், பயன்பாடுகளையும், விதிமுறைகளையும் விரிவான அளவில் தெரிந்து கொள்ள முடிகிறது.

(iv) அடுத்தபடி, மிக முக்கியமாக, மாதிரி அளவை மதிப்புக்கும், முழுமைத்தொகுதி அளவைக்கும் இடையிலோ, அல்லது ஒரு உண்மையான அளவையின் இரு மதிப்பீடுகளின் இடை

யிலோ, உள்ள வேறுபாட்டின் அளவினை 'மிகைத்தன்மை சோதனை' மூலம் நன்கு அறிய இயலுகிறது (Test of significance.).

(v) இதன்படியாக எந்த ஒரு முழுமைத் தொகுதி அளவைக்கும் (parameter) அதனது 'நம்பக எல்லை'யினை (confidence limits) வரையறை செய்ய முடிகிறது.

(vi) ஒரு மாதிரி மதிப்பீட்டின் நம்பகத் தன்மையினையும் (reliability) திட்பத்தினையும் (precision) நன்கு உறுதி செய்ய இயலுகிறது.

vii) எடுக்கப்படும் மாதிரியின் உத்தம அளவினை (optimum size) வரையறுக்க முடிகிறது. முடிவுகள் சிறப்பானதாக அமைய வழி ஏற்படுகிறது.

துணை மாதிரிகள் (subsamples) நன்கு ஆராயப்பட்டு வகைப் படுத்தப்படுகின்றன. இவற்றின் அடிப்படையில் பெறப்பட்ட முடிவுகள் ஒப்பிடப்பட்டு ஆய்வு நடைபெறும் நிலையினை அறிய வழி ஏற்படுகிறது.

ஒரு சிறந்த மாதிரித்திட்ட அமைப்பை உருவாக்குவதென்பது கள ஆய்வுப்பணியை நன்கு அமைப்பதிலும் அட்டவணைப் படுத்தும் நிலையில் உள்ள திட்டவரையறையினை அமைப்பதிலும் அடங்கியுள்ளது.

மேலும் அதிக அளவில் தரமான விவரங்களைப் பெறுவதும் அட்டவணைப்படுத்தப்பட்ட நிலையில் நன்கு திட்டமிட்டு இடர்ப் பாடுகளைக் களைவதும் இன்றியமையாதது.

8. முன்னோடி ஆய்வு (Pilot Survey)

ஒரு ஆய்வினை திறம்படத் திட்டமிடுவதானது முழுமைத் தொகுதி, நடைமுறை இடர்ப்பாடுகள், ஆய்வுச் செலவு இவற் றைப்பற்றிய சில முன்னோடி (prior) விவரங்களைச் சேகரிக்கும் நிலையிலும் முடிவுகளை அட்டவணைப்படுத்தும் கட்டத்திலும் பெறுவதைப் பொறுத்து அமைகிறது. பழைய ஆய்வுகளினின்று இவ் விவரங்கள் கிடைக்கப்பெறாத நிலையில் ஒரு முன்னோடி (அல்லது முன் நிலை) ஆய்வினைத்திட்டமிட்டு நடத்துவது தேவை யாகிறது. இதன்மூலமாக ஆராயப்படும் இயல்புகளின் வேறு பாடு (variability) பற்றிய முன்னோடி விவரங்களும் பலவகைத் திட்டங்களில் விவர சேகரிப்பு மற்றும் அட்டவணைப்படுத்தல் இவற்றுக்கு ஆகும் செலவினையும் பற்றிய விவரமும் கிடைக்

கின்றன. இதன்மூலம் முக்கிய ஆய்வுத் திட்டமிடலில் செலவு மற்றும் வேறுபாடு இவற்றின் சார்புகளை (functions) உருவாக்க இயலும். மேலும் இத்தகைய முன்னோடி ஆய்வானது (i) உருவாக்கிய வினைப்பட்டியல், கொடுக்கப்படும் விளக்கங்கள் ஆகியவற்றைச் சோதனை செய்தறியவும், (ii) கள ஆய்வு, அட்டவணைப்படுத்தல் இவற்றுக்குப் பொருத்தமான வழிமுறைகளைத் தேர்ந்தெடுக்கவும், (iii) பணியாளரை (staff) நன்கு பயிற்சி பெறச் செய்யவும் உதவிபுரிகிறது. ஆய்வின் நோக்கத்தைப் பொறுத்து இது அமையும்.

9. களப் பணி (Field Work)

ஆட்சித் துறையில் ஏற்கெனவே இருக்கும் பணியாளரைப் பயன்படுத்துவதன்முலம் ஆய்வுச் செலவு மிகவும் குறைவாக அமையலாம். விவரங்கள் எளிதாக இருந்திடின் (simple), இம்முறை எதிர்பார்க்கும் அளவிற்குப் பயன்தரும். வேறொரு முறையானது, ஆய்வுப் பணிக்கேற்ற கள ஆய்வாளர்கள் அடங்கிய ஒரு நிரந்தர அமைப்பினை உருவாக்கி, அவர்களுக்குச் சிறந்த பயிற்சியையும் அளிப்பதாகும். முந்திய ஆய்வுகளில் பெறும் அனுபவங்கள், பின்னால் தொடரும் பல ஆய்வுகளில் பயன்படும். மேலும் ஆய்வு மேற்பார்வை (supervision), பரிசீலனை (scrutiny) போன்றவை சிறந்து விளங்கும்.

கள ஆய்வில் அமையும் ஆய்வாளர் பயிற்சி, மேற்பார்வை போன்ற விவரங்களை, நடைமுறைக் கண்ணோட்டத்துடன் நன்கு அடுத்துகொள்ளுதல் அவசியம் இவற்றைப் பொறுத்து, சேகரிக்கப்படும் விவரங்களின் தன்மைக்கு ஏற்ப, ஆய்வு முடிவுகளின் (survey results) இயல்பு அமைந்திருக்கும். பல்வேறு சொற்கள், கருத்துக்கள் பற்றி அறிந்திருப்பதோடல்லாமல், பலவழிகளில் (sources) சரியான விவரங்களைப் பெறக்கூடிய திறனும் ஆய்வாளரிடம் அமையவேண்டும். பயிற்சி அளித்தற்கான கட்டங்கள் (stages) குறைந்தனவாக இருக்கவேண்டும்.

பரிசோதனை முறையானது (inspection) ஆய்வுப் பணியினைப் பலவகைகளில் சரிபார்த்தலை (check-up) கொண்டதாகும். களத்திலேயே அந்தந்த நிலைகளில் நேரடிச் சோதனை (on the spot check) நடத்தலாம். பூர்த்தி செய்யப்பட்ட படிவங்களை (forms) நன்கு பரிசீலிக்கலாம். ஆய்ந்தறியப்பட்ட விவரங்களின் ஒரு தெரிவு செய்யப்பட்ட பகுதியைக் (random selection) கருதி, அவற்றின் தன்மையை மதிப்பிடலாம். வடிவமைப்பு, மற்றும் பல்வகைச் செயல்முறை (designing & processing) பணி

யாளருடன், மேற்பார்வை அதிகாரிகள் தொடர்ந்து தொடர்பு பெற்றிருப்பின், ஆய்வில் பல்வேறு நிலைகளில் நிகழும் எதிர்பாராத அல்லது இயல்புக்கு மாறான சிக்கல்களைத் தீர்ப்பது எளிதாகும்.

பலவகைப் பணிகள் நடத்தப்பெறும் கால அளவுகளைக் குறித்துவைப்பது அவசியமானது. மாதிரி 'சட்ட' வடிவமைப்பு, கூறுகளைத் தேர்வு செய்து தொடர்பு கொள்வது, ஆய்வு போன்றவை நடைபெறும் கால அளவைப் பற்றிய குறிப்பினை (record) வைத்தல் நலம்.

10. ஆய்வு விவரங்களின் பகுப்பாய்வு (Analysis of Survey Data)

மூன்று நிலைகளில் இத்தகைய பகுப்பாய்வு (Processing) நடைபெறும்.

அட்டவணைப்படுத்தல் (Tabulation)

பெரிய அளவில் கண்டறிந்து குறிக்கப்பட்ட விவரங்களை, அர்த்தமுள்ள விரிவான அட்டவணைகளாகச் சுருக்குதல் முக்கியம். ஆய்வின் அடிப்படை நோக்கம் மாதிரிச் செயல்முறை, சேகரித்த முறை, முடிவுகளில் அனுமதிக்கக் கூடிய நிர்ணயித்த பிழை எல்லை (margin of error) ஆகிய விவரங்களுடன் அட்ட வணைகள் தரப்படவேண்டும். இதன்மூலம், கொடுக்கப்படும் விவரங்களைத் தவறாகப் பயன்படுத்துவது தவிர்க்கப்படுகிறது.

பொருள் பற்றி விரிவாய்வு (Subject Analysis)

சுருங்கிய அளவில் அமையும் அட்டவணைகள் மூலம், அவற்றில் உள்ள பல்வேறு இயல்புகள் (Characteristics) பற்றிய தொடர்புகள் (relationslhips), போக்குகள் (trends) போன்ற வற்றை அறியவேண்டியும் மற்றவழிகளில் (sources) பெற்ற விவரங்களுடன் ஒப்பிடவேண்டியும், பலவகைப்பட்ட குறுக்கீட்டு (cross) அட்டவணைகள் மூலம் அமைத்து விரிவான அளவில் சிறப்புப் பகுப்பாய்வு நடத்தப்பெறும். இதற்கான கால அளவும் அதிகமாகும்.

புள்ளியியல் பகுப்பாய்வு (Statistical Analysis)

பின்னால் நடைபெறும் ஆய்வுகளின் திறத்தை அதிகப் படுத்துவதற்கேற்ப, செயல்முறையினைச் செம்மைப்படுத்தும் படியான முறைகளைத் தீர்மானிக்க மேற்கொள்ளும் பகுப்பாய்

வாகும். மாறுபாடு, மற்றும் செலவுபற்றிய (variance & cost) விரிவான ஆய்வும், பல நிலைகளில் நிகழும் மாதிரி சாராப் (non sampling) பிழைகளை நன்கு ஆராய்வதும் இதில் அடங்கும்.

11. செயல்முறைப்படுத்துதல் (Processing)

ஆய்வு விவரங்களைத் தக்கவாறு தரம் பிரித்து, வகைப்படுத்தி (sorting), நன்கு பரிசீலித்தபின்னர், மனிதமுயற்சியால் அல்லது பொறிவழிமூலம் (manual & mechanical) பலவகை மதிப்பீடு களைக் கணக்கிட்டு (computation), அட்டவணைகளைப் பெற்று முடிவுகளை (results) வரையறுப்பது செயல்முறைப் (process) படுத்தும் கட்டமாகும். கணக்கீடுகளை (Computation) உடனுக் குடன் சரிபார்த்துத் திருத்தியமைப்பதன்மூலம் (correction), மாதிரி சாராத (non-sampling) பிழைகளைக் குறைக்க முடிகிறது. பல நிலைகளிலும் தடங்கல் அல்லது சிக்கல் ஏற்படா வண்ணம் பணி நடைபெற, விரிவான அளவில் திட்டமிடுதல் (planning) வேண்டற்பாலதாகும்.

மிகவும் பாதிக்கக்கூடிய குறைபாடுகளை (defects) எவ் வகையிலேனும் கண்டறிந்து அவற்றை நீக்க முயலுதல் இன்றியமையாதது. விபரங்கள் பன்முறை நன்கு பரிசீலித்தபின் (scrutiny) ஏற்றுக்கொள்ளப்படுதல் நன்று. விவரங்களின் தன்மைக்கு ஏற்றபடி, தேவைப்படும் அட்டவணைபற்றிய முடிவுகளை முன்கூட்டியே வகுத்தல் வேண்டும். அவற்றோடு, மாதிரிகளின் அளவுகள் (sizes), ஆய்வுப் பணிகளில் ஏற்பட்ட செலவு (cost of survey), பிழையின் எல்லை அளவு (margin of error) போன்ற விவரங்களையும் தக்க பட்டியல்கள்மூலம் வெளிப்படுத்துதல் அவசியம்.

மிகப் பெரிய அளவில் (large-scale) நடத்தப்படும் அள வெடுப்புகளில் (surveys), மின்விசை (electrical or electronic) மூலம் செயல்படும் பொறிகளைப் (equipments) பயன்படுத்த வேண்டும். (உதாரணம்: மின்விசைக் கணக்கீட்டுப் பொறி—electrical calculator) விவரங்களைப் பொருத்தமானபடி. எண் குறியீடுகளால் (code) குறிப்பிட்டு, அவைகள் அனைத்தையும் துளை அட்டைகள் (punch cards) அல்லது வேறு சாதனங்களில் தொகுத்து வைப்பதன்மூலம், செயற்படுத்துதலும் (processing) பல்வேறு கணக்கீடுகளைப் பெறுவதும் (computation) எளிதாக இருக்கும். மாறாக, பயிற்சிபெற்ற நபர்கள்மூலம் (manual) செயல்முறை அமைந்திருப்பின் (processing), தேவையான குறிப்புகளும் (instructions) விளக்கங்களும் கொடுக்கப்பட்டு இருத்தல் இன்றியமையாதது.

கணக்கிடப்பட்ட மதிப்புகள் துல்லியமாக (accurate) அமைதல் மிகஅவசியம். அவற்றைப் பொறுத்தே இறுதிமுடிவுகள் எடுக்கப்படுகின்றன. கணிப்பில் (numerically) தவறுநிகழ்ந்தால், அடிப்படை நோக்கமே பாழாக வழியுண்டு. திருப்திகரமான அளவிற்கு, கணக்கீடுகள் ஒரு முறைக்கு இருமுறையாகச் சரி பார்க்கப்படுதல் நன்று.

செயல்முறைப் (process) பணிகளில் ஏற்படும் கால அளவை (time period) குறித்துக்கொள்வதன்மூலம், (recording), பல்வேறு முறை பற்றிய நன்மைகள் (advantages), பணிபுரிவோரின் திறமை (efficiency), வகைப்படுத்த ஏற்படும் செலவு (cost). போன்றவை பற்றிய ஒரு மதிப்பீடு கிடைக்க வழி உண்டு. பிற்காலங்களில் நடத்தப்பெறும் ஆய்வுகளில் இது மிக உதவுகிறது. ஆய்வுச் செலவு சார்பையும் (cost function), காலக் குறிப்பையும் (time-recording) பற்றிய ஆராய்ச்சியை மஹலநோபிஸ் (Mahalanobis) 1944-ல் விரிவாக நடத்தியிருக்கிறார்.

12 இறுதி அறிக்கை தயாரித்தல் (Final Reports)

மாதிரி அளவெடுப்புகளில் பெறப்படும் இறுதி முடிவுகள் (results) பற்றிய விரிவான அறிக்கையினைத் (report) தயார் செய்கையில், பின்வரும் விவரங்களைக் கருத்தில் கொள்ள வேண்டும் :—

(1) ஆய்வின் அடிப்படை நோக்கமும் எல்லை வரையறையும் (Objective & Scope)

ஆய்வு நோக்கமும், முடிவுகளைப் பயன்படுத்தும் விதமும் தெளிவாக்கப்பட்டு, ஆய்வுக்கு எடுத்துக் கொள்ளப்பட்ட எல்லைப் பரப்பு, முழுமைத் தொகுதி இவைபற்றிய விரிவான விளக்கமும் தரப்பட வேண்டும்.

(2) பொருள் (Subject) பற்றிய தெளிவு

சேகரிக்கப்பட்ட விவரக் கூறுகளை விரிவாக வரைதல் வேண்டும். அவற்றைப் பற்றிய ஒரு மாதிரிப் பட்டியலை (specimen Schedule) இணைத்தல் நன்று.

(3) விவரச் சேகரிப்பு முறை

கண்டறியும் நிலைமையில் ஏற்பட்ட நடைமுறை சிக்கல்கள், கடினங்கள் பற்றி விரிவாக்கப்படல் வேண்டும்.

(4) மேற்கோள் மற்றும் அறிவிப்புக் கால அளவு: (Reference and Reporting Periods)

ஆய்வுக் காலத்தை (Survey Period) தெளிவாகக் குறிப்பிட்டு, விவரங்கள் சேகரிக்கப்படும் காலத்தைத் தெளிவாக வரைதல் அவசியம்.

(5) மாதிரி வடிவமைப்பு, மதிப்பிடும் முறை (Sample Design & Estimation method)

மாதிரிக் கூறுகள் பற்றி விரிவாகக் குறிப்பிட்டுக் கையாண்ட மாதிரி முறையினை (Sampling method) விளக்குவதோடு, மாதிரியின் அளவு (Size), அந்த அளவினைத் தீர்மானிக்கும் அடிப்படை போன்றவற்றையும் திட்டவட்டமாக குறிப்பிடவேண்டும். மதிப்பீட்டு சூத்திரங்களை (Estimation formulae) நன்கு குறிப்பிட வேண்டும்.

(6) அட்டவணைப்படுத்தல், முடிவுகளை அளித்தல் (Tabulation & Presentation of Results)

பட்டியல் அமைக்கக் கையாளும் முறைகள், இதில் உள்ள பல்வேறு கட்டங்கள் (stages), மற்றும் பல கட்டங்களில் ஏற்படும் இன்னல்கள் (Problems) அவற்றிலிருந்து விடுபடும் விதம் போன்ற விவரங்கள் குறிக்கப்படவேண்டும்; தெளிவான (unambigous) தலைப்புகளுடன் (titles), விரிவான பத்தித் தலைப்புகளுடன் அட்டவணைகளை அளிப்பது, ஆராய்ச்சிக்கு எளிது. இதற்குத்துணை செய்யும் விளக்கப் படங்கள் (charts), வரைபடம் (graphs) ஆகியவற்றையும் உருவாக்கலாம்.

(7) துல்லியத்தின் அளவு (Measurement of Accuracy)

முடிவுகளில் உள்ள துல்லியத்தின் அளவினைக் குறிப்பிட வேண்டும். விவரங்களின் துல்லியம் பற்றி அறியச் செய்யப்படும் ஒப்புமைகள் (camparisons) சோதனைகள் (tests) பற்றியும் நன்கு விளக்குதல்வேண்டும். ஆய்வில் காணப்படும் பதிலின்மை (non-response) பற்றிய ஒரு கருத்தும் தெளிவாக்கப்படுதல் ஒன்று.

(8) செலவு பற்றிய பகுப்பு (Cost Structure)

துவக்கப்பணி (Preliminary work). மாதிரித் தேர்வு, களப்பணி (field work) வகைப் படுத்தல், (process) போன்ற பல கட்

டங்களில் (Stages) ஏற்படும் செலவுபற்றிய விவரங்களைத் தர வேண்டும். ஆய்வின் மொத்தச் செலவும் மதிப்பிடப் படல் நன்று.

(9) பொறுப்பின் வரையறை (Responsibility)

ஆய்வினை (investigation) திட்டமிட்டுச் செயல்படுத்தும் நிறுவன அமைப்பு (organizational set-up)பற்றி விரிவாகக் குறிக் கப்படல் தேவையான ஒன்று.

(10) மேற்பார்வை பற்றிய குறிப்பு (Reference)

ஆய்வுப் பணியில் பயன்படுத்தப்பெறும் தேவையான பதிப் பிட்ட வெளியீடுகள் (published material) மற்றும் அறிக்கைகள் (Reports) ஆகியவைபற்றிய தெளிவான குறிப்புகளைத் தருவது அவசியம். தேவைப்படும்போது அவற்றை மேற்பார்வைக்குக் கருதலாம்.

(13) மாதிரி அளவினை மதிப்பிடல்: (Estimation of Sample Size)

(அ) ஒரு மாதிரியின் அளவினை (size) தெரிவு செய்யப் பின்வரும் முக்கிய நிலைகளைக் (steps) கருதுகிறோம்:

(1) அந்த மாதிரியிலிருந்து எதிர்பார்க்கப்படும் ஒரு முடிவினைப்பற்றி ஒரு விரிவான அறிக்கை (statement) அமைக்கப் பட்டிருக்க வேண்டும்; பிழைகளின் அனுமதிப்பு எல்லை வரையறை (desired limits of errors) மூலமாகவோ, அல்லது அறிந்த மாதிரி முறை பற்றிய முடிவுகளிலிருந்து எடுக்கப்படும் நடவடிக்கை மூலமாகவோ, இத்தகைய அறிக்கை தரப்படலாம்.

(2) மாதிரி அளவு எண்ணுள 'n'-ஐ, அதன் தேவைப்பட்ட திட்ப (precision) அளவுடன் தொடர்புபடுத்தும் ஒரு சமன்பாட் டைக் காணவேண்டும். இத்தகைய சமன்பாடானது, (equation) திட்ப அளவு பற்றிய அறிவிப்பினையும், கையாளப்படும் மாதிரி முறையினையும் பொறுத்து மாறுபடும். வாய்ப்பின் அடிப்படையில் அமைந்த மாதிரி முறைகளில் (probability sampling), இத்தகை யசமன்பாடு எளிதில் அமைக்கப்படுவது குறிப்பிடத்தக்கது.

(3) முழுமைத் தொகுதியின் சில அறியாத (unknown) இயல்பு களை, மாறிலி அளவைகளாக (parameters) மேற்குறிப்பிட்ட சமன்பாடானது கொண்டு விளங்கும். இத்தகைய அளவைகளை மதிப்பிட்டு அறிவதன் மூலமாக. குறிப்பிட்ட முடிவுகளைக் காண் பது எளிதாகிறது.

(4) பொதுவாக, ஒரு முழுமைத்தொகுதியினுடைய சிலமுக்கிய உட்பிரிவுகட்கு முடிவுகளை வெளியிடவேண்டியிருப்பதால், ஒவ்வொரு உட்பிரிவிலும் தக்கபடியான பிழைகளின் எல்லை வரையறை கொடுக்கப்படும். இத்தகைய நிலையினுல் ஒவ்வொரு உட்பிரிவிற்கும் ஏற்ற அளவில் ஒரு மாதிரி அளவு (size) *n* ஆனது கணக்கிடப்பட்டு, இறுதியில் அவற்றின் கூட்டு மதிப்பானது பெறப்படும்.

(5) ஒரு மாதிரி முறை அளவில், ஒன்றுக்கு மேற்பட்ட இயல்புகள் (charctaeristics) அளவிட்டுப் பெறப்படுகின்றன. சில நேரங்களில், எடுத்துக்கொள்ளப்பட்ட கூறுகள் அதிக அளவில் (large) அமைய நேரிடுகையில் ஒவ்வொரு கூறினுக்கு அல்லது இயல்பினுக்கு ஏற்ற திட்ப அளவு (precision) குறிப்பிடப்பட்டு. பெறப்படும் *n*- மதிப்புக்கள் சிக்கலானவையாக அமைகின்றன. எனவே, இவற்றை எளிதாகக் காணும் வகையில் செயல்முறை ஒதுக்கப்படவேண்டும்.

(6) அடுத்து, தெரிவு செய்யப்படும் மாதிரி அளவு *n*- ஆனது மாதிரி எடுத்தற்குரிய சாதனங்களின் தன்மைக்கும் பொருத்தமாக (consistent) அமைதல் அவசியம். அதாவது, கருத விழையும் (proposed) ஒரு மாதிரி அளவினைப் பெற இயலும் வண்ணம், செலவு, தேவைப்படும் கால அளவு, கிடைத்திடும் விவரங்கள் ஆகியவைபற்றி தெளிவாக மதிப்பிட்டு அறிதல் வேண்டும். சில சமயம் *n*-ன் அளவு மிகக் குறைக்கப்படுவது தவிர்க்க இயலாதது. அத்தகைய நிலையில் குறைந்த அளவுடைய (small-size) மாதிரியினைக் கருதுதல் தேவையா என முடிவு செய்து கொள்ள நேரிடும். இதன் மூலம் திட்ப அளவு மிகவும் குறைய நேரிடுவதால், தக்க சாதனங்கள் (material) கிடைத்திடும் வரை எடுக்கப்பட்ட முயற்சிகளைக் கைவிடவும் செய்யலாம்.

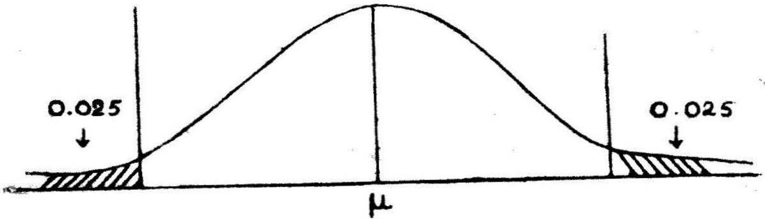
இனி, மாதிரி அளவானது எந்த ஒரு மாதிரிக் கூறினுக்கும், திட்ப அளவுடன் பெற்றிருக்கும் தொடர்பினைச் சற்று விரிவான முறையில் ஆய்வு செய்கிறோம்.

(ஆ) திட்ப அளவு: (Precision) ஒரு கணக்கீடு

மாதிரி முறைப் பிழையானது, t-என்ற மதிப்பீட்டிற்கு, t—E (t) என்று அமையும். முழுமைத் தொகுதி அளவையினை (Parameter), மாதிரிமுறை அளவையின்மூலம் குறிப்பிடுகையில் θ என்ற முழுமை அளவைக்கு ஒரு நம்பக இடைவெளியினை (Confidence interval) காணலாம். ஒரு குறிப்பிட்ட நம்பகத்

தன்மை அளவிற்கேற்ப (reliability) திட்பத்தினைக் கணக்கிடலாம். அதாவது θ எனும் முழுமைத் தொகுதி அளவையானது ஒரு குறிப்பிட்ட எல்லை வரையறைக்குள் (limits) அமைவதாகப் பெறுகிறோம்.

இயல்நிலைப் பரவல் (normal distribution) தன்மை மூலமாக, 95% என்ற நம்பகத்தன்மை அளவிற்கு ஏற்ற திட்ப அளவினைக் காண்போம். கருதப்பட்ட அளவை, μ என்ற முழுமைத் தொகுதி சராசரியாகும்.



படம் : 3

இயல்நிலை (normal law) விதிப்படி,

$$P\left\{-1.96 < \frac{\bar{x}-\mu}{\sigma_{\bar{x}}} < 1{\cdot}96\right\} = 0{\cdot}95$$

$\bar{x}$ → மாதிரிச் சராசரி மதிப்பு [$E(\bar{x}) = \mu$]

$\sigma_{\bar{x}}$ → அச் சராசரியின் மாறுபாடு

இதிலிருந்து

$P(-1{\cdot}96\sigma_{\bar{x}} < \bar{x} - \mu < 1{\cdot}96\,\sigma_{\bar{x}}) = 0{\cdot}95$ என்பது, $\bar{x}-\mu$ என்னும் இடைவெளியானது கொடுத்த வரம்பில் (interval அமைவதற்கான வாய்ப்பு 0·95 என்பதான நிலையைக்குறிப்பிடும். அல்லது $|\bar{x}-\mu|$ என்ற இடைவெளி மதிப்பு $1{\cdot}96\,\sigma_{\bar{x}}$-ஐவிடக் குறைந்து விளங்கும் வாய்ப்பு 0·95 ஆக இருக்கும் என காட்டுகிறது.

இதிலிருந்து, 95% நம்பக அளவில் (confidence) μ என்னும் சராசரிக்கு மதிப்பிடப்பட்ட நம்பக எல்லை (confidence limit). $\bar{x} \pm 1{\cdot}96\,\sigma_{\bar{x}}$ என்பதாகும். மேற்கண்டவற்றிலிருந்து, $|\bar{x}-\mu|$ என்ற இடைவெளியின் (வேறுபாட்டின்) (difference) உத்தம அல்லது உச்ச (maximum) அளவானது $1{\cdot}96\,\sigma_{\bar{x}}$ என்று பெறப்படும். இந்த வேறுபாடு மதிப்பு (difference) ஒரு கொடுக்கப்பட்ட மாதிரிக்கு மட்டும் பொருந்துவதல்ல. மாதிரிப் பிழையிலிருந்து இது வேறுபடும்.

கொடுக்கப்பட்ட ஒரு மாதிரி அளவுடன் (size) மற்றும் நம்பகத் தன்மையுடனும் (reliability), பல மாதிரிகளைக் கருத்திற் கொண்டு பெறப்பட்ட, ஒரு மதிப்பீட்டிற்கும் (estimator), முழுமைத் தொகுதி அளவைக்கும் (parameter) இடையே உள்ள உச்ச அளவு வேறுபாடு திட்பஅளவு (precision) அல்லது நுட்ப அளவு என்று கூறப்படும். திட்ப அளவை கணக்கிட, முழுமைத் தொகுதி அளவை (population parameters) தெரிந்திருக்க வேண்டிய அவசியமில்லை. மாறுபாடு (variance) மட்டும் தெரிந்திருத்தல் வேண்டும்.

எடுத்துக்காட்டு :

ஒரு மாதிரியின் சராசரி ($\bar{x}$) மதிப்பு 2·5 என்றும், அச் சராசரியின் மாறுபாடு ($\sigma_{\bar{x}}$) 1·67 என்றும் காணப்பட்டு, $n = 2$ எனவும், நம்பகத்தன்மை 90% எனவும் கொடுக்கப்படின் நம்பக எல்லையினை, (Confidence interval)

$$\bar{x} - 1{\cdot}64\,\sigma_{\bar{x}} < \mu < \bar{x} + 1{\cdot}64\,\sigma_{\bar{x}}$$

என்று முழுமைத் தொகுதி சராசரி μ - க்கு வரையறுக்கலாம்.

அதாவது

$2{\cdot}5 - (1{\cdot}64 \times 1{\cdot}67) < \mu < 2{\cdot}5 + (1{\cdot}64 \times 1{\cdot}67)$ அதாவது $0{\cdot}4 < \mu < 4{\cdot}6$ என்பதாகும்.

இதில்உள்ள $\frac{1}{2}$x (நம்பக இடைவெளி, அதாவது நம்பகஇடைவெளியின் பாதிமதிப்பு $1{\cdot}64 \times 1.67$ என்பது அதாவது 2.1 ஆனது திட்ப அளவாக கருதப்படும்.

ஒரு மதிப்பீட்டின் பிழை வரம்பினை (margin of error), அனுமதிப்புப் பிழை என்று கூறுகிறோம் (permisible error). உதாரணம் : $|\bar{y} - \bar{Y}|$ என்ற அளவு.

(இ) மாதிரி அளவும் திட்பமும் (Sample size and Precision)

நம்பகக் கெழுக்கள் (Confidence) நம்பகத் தன்மையின் (reliability) அளவைக் குறிப்பிடுவன. இதனைப் பல நிலைகளில் $(1-\alpha)$ என்று குறியிடுவோம். $(1-\alpha) = 0{\cdot}90$; $(1-\alpha) = 0{\cdot}95$ என்பவை போன்று அமையும் அளவீடுகள், 90%, 95% என்ற அளவில் நம்பகத் தன்மையினைச் சுட்டிக் காட்டும். ($\alpha = 0{\cdot}05$ என்பது, சரியான முடிவிலிருந்து தவறுவதற்கு ஏற்படும் பிழை வாய்ப்பினை (5%) குறிக்கும்)

∴ ஒரு மதிப்பீட்டிற்கு, (error) திட்ப அளவு, d = (நம்பகத் தன்மை) × (தரப்பிழை) என்றபடி வரையறை செய்யலாம்.

நம்பகத்தன்மை, (reliability), மற்றும் தரப்பிழை அளவு (standard error) இவற்றின் தொடர் விளைவு (Product) திட்ட அளவை மதிப்பிடும்.

உதாரணமாக, மாதிரி முறையில் $\bar{y}$ என்ற மாதிரிச் சராசரி மதிப்பீட்டின் திட்ப அளவினைக் குறிப்பிட்ட நம்பகத்தன்மை அளவில் கணக்கிடலாம். 95% என்பது நம்பகத் தன்மை எனில், அதற்கேற்ற மதிப்பு 1·96 என்பதை பயன்படுத்தி, கொடுக்கப் பட்ட தரப் பிழை அளவையும் ($\sigma_{\bar{y}}$) கருதி,

$$|\bar{y} - \bar{Y}| = 1.96\,\sigma_{\bar{y}}$$

என்று திட்பஅளவு பெறப்படும். $\bar{Y}$: முழுமைத் தொகுதி அளவை.

பொதுவாக, நம்பகத்தன்மை கெழு (reliability coefficient) மதிப்பு z எனில் $\bar{y}$-ன் திட்பஅளவு $= z\sigma_{\bar{y}}$ என்று ஆகும்.

மாதிரியின் அளவு அதிகரிக்கப்படும்போது, திட்பம் அதிகரிக்கிறது. அதாவது $z\sigma_{\bar{y}}$ என்ற மதிப்பு குறையும்; அதற்கேற்ப மதிப்பீட்டு திட்பம் அதிகமாகும்.

வரையறைப்படி $d = |\bar{y} - \bar{Y}| = z\,\sigma_{\bar{y}}$

$\bar{y}$-ன் தரப்பிழையானது $\frac{\sigma}{\sqrt{n}}$ ஆகும். σ^2 என்பது முழுமைத் தொகுதி மாறுபாடு. n- என்பது மாதிரியின் அளவு. 95% நம்பகத் தன்மைக் கேற்ப, $z = 1{\cdot}96$ ஆகும். முழுமைத் தொகுதி மாறு பட்டு மதிப்பு (σ^2) 400, அதாவது $\sigma = 20$ எனக் கொள்க.

அப்போது, $d = z.\frac{20}{\sqrt{n}}$

$$= \frac{1{\cdot}96 \times 20}{\sqrt{n}} \text{ ஆகும்.}$$

மாதிரி அளவு $n = 16$ என்றால்,

$$d = \frac{1{\cdot}96 \times 20}{\sqrt{16}} = 9{\cdot}8 \text{ ஆக அமையும்.}$$

மாதிரி அளவு $n = 25$ என்றால்,

$$d = \frac{1{\cdot}96 \times 20}{\sqrt{25}} \quad 7{\cdot}84 \text{ ஆகும்.}$$

மாதிரி அளவு $n = 36$ எனில்,

$$d = \frac{1{\cdot}96 \times 20}{\sqrt{36}} = \frac{39{\cdot}2}{6} = 6{\cdot}54$$

எனவே, n அதிகமாகும்போது, d குறைய மதிப்பீட்டுத் திட்பம் அதிகமாகிறது. அதாவது, மாதிரி அளவு அதிகரிக்கப்பட, மதிப்பீடானது முழுமைத் தொகுதி அளவைக்கு மிக அருகில் இருக்கும் தன்மை பெறுகிறது எனலாம்.

கொடுக்கப்பட்ட திட்பத்திற்கு ஏற்ற மாதிரி அளவினை (n) கண்டுபிடிக்கலாம்.

உதாரணமாக, 95% நம்பகத் தன்மையில், $\bar{y}$ என்பதின் திட்ப அளவு 4 என்னும் மதிப்புக்கு உட்பட்டதாக அமையத்தேவைப் படும் மாதிரி அளவு n, பின்வருமாறு கணிக்கப்படும் :

$d = |\bar{y} - \bar{Y}| \leq 4$ (கொடுக்கப்பட்டது.

அதாவது, $d = z.\ \dfrac{\sigma}{\sqrt{n}} = \dfrac{1.96 \times 20}{\sqrt{n}}$ $(\sigma = 20)$

இதிலிருந்து பெறும் $n = 10^2 = 100$ (ஏறத்தாழ) என்று கணக்கிடப்படுகிறது.

[அதாவது $n = \left(\dfrac{1.96 \times 20}{4}\right)^2 \doteq 100$]

எடுத்துக்காட்டு: (1) 1000 மாணவர்களது சராசரி எடையினை 5 *lb* (பவுண்டு) அளவிலான திட்பத்துடன், 95% நம்பகத் தன்மைக்கேற்ப மதிப்பீடு செய்ய மாதிரி அளவு என்ன ?

மாதிரியின் அளவு n ஆனது கண்டுபிடிக்கப்படவேண்டும்.

$d = 5$; $z = 1.96$

$\sigma = 15$ *lb* என்று எடுத்துக்கொள்வோம். (கொடுக்கப் பட்டிருக்கவேண்டும்.)

எனவே $n = \dfrac{(1{\cdot}96 \times 15)^2}{5^2}$ ஆகும்

அதாவது $n = 34{\cdot}6 \doteq 35$ (திருத்தமாக)

மேலும், 99·7% என்ற நம்பக வீதத்துக்கு ஏற்ப $z = 3$ என்ற மதிப்பு அமையும்.

$$\therefore\ n = \left(\frac{3 \times 15}{5}\right)^2 = 9^2 = 81 \text{ ஆகும்.}$$

எடுத்துக்காட்டு (2) :

ஒரு பெரிய தொகுதியிலிருந்து மாணவர்களின் சராசரி எடை

மதிப்பிடப்பட வேண்டும். 3 *lb* அளவு திட்டமும், 2 σ அளவு நம்பகத் தன்மையும் கொண்டு, எடுக்கப்படும் மாதிரியின் அளவு என்ன ? $\sigma = 12$ *lb* என்று கொள்க.

$$\text{வரையறை:} \quad n = \left(\frac{z.\sigma}{d}\right)^2$$

$$z\sigma = 2 \times 12 = 24.$$

$$d = 3.$$

$$\text{எனவே} \quad n = \left(\frac{24}{3}\right)^2 = 8^2 = 64 \text{ ஆகும்.}$$

64 எண்ணிக்கை அளவுள்ள மாதிரி எடுக்கப்படவேண்டும்.

குறிப்பு : பின்பு கருதப்படும் வெவ்வேறு மாதிரி முறைகள் பற்றிய விரிவாய்வுகளில், ஒவ்வொரு முறையின் இறுதியிலும் தனித் தனியாக மாதிரி அளவு நிர்ணயிக்கப் பெறும் விதத்தினை விளக்குகிறோம்.

(ஈ) பல்வேறு இயல்புகள் கருதப்படுகையில். மாதிரி அளவு நிர்ணயித்தல் (Sample Size with more than one Item)

பெரும்பான்மை ஆய்வுகளில், பல இயல்புகள்பற்றி விவரங்கள் பெறவேண்டியதாக இருக்கும் சூழ்நிலை ஏற்படுகிறது. அவற்றுள் மிக இன்றியமையாத இயல்புகளை மட்டும் கருதிய அளவில் பிழைஎல்லைகளை (Margin of error)வரையறுத்து,மாதிரி அளவினை நிர்ணயித்தல் அவசியமாகிறது. முதலில் இத்தகைய முக்கியமான இயல்புகள் ஒவ்வொன்றுக்கும் தனித்தனியாக மாதிரி அளவினைத் தீர்மானிக்கவேண்டும். இத்தகைய ஒரு இயல்பு கருதிய *n*-ன் மதிப்பீடுகள் பெறப்பட்டபின், நிலையினை முழுதும் கருத்திற்கொண்டு பின்வருமாறு காணலாம்:

தேவைப்படும் வெவ்வேறு *n*-மதிப்புகள் ஏறக்குறைய சம அளவில் இருப்பின், கருதப்பட்ட செலவு நிலைக்கு ஏற்ப அமையும் பெரிய அளவினதான *n*-ஆனது. தெரிவு செய்யப்படும். *n* -மதிப்புகள் தம்முள் அதிக அளவில் வேறுபடும் நிலையில், பெரிய அளவினைத் தெரிவு செய்தல் இயல்பாக அமையும். ஏனெனில் அதன்மூலம் அதிக அளவு திட்பம் அமைந்திருப்ப தாகக் கருதப்பட வாய்ப்பு உண்டு. எனவே, குறைந்த *n* அளவு களையும் பயன்படுத்தத் தக்க விதத்தில் , திட்ப அளவினை (Standard of precision) சில இயல்புகட்கேனும் தளர்த்துவது மேற்கொள்ளப்படலாம்.

சில நிலைகளில் n மதிப்புக்கள் வெவ்வேறு இயல்புகட்கும் ஒழுங்கற்ற முறையில் மாறுபடும் தன்மை பெற்றிருப்பின் சரியான திட்ப அளவு நிர்ணயிப்பது இயலாத நிலையில், சிலவற்றைக் கருதாமல் விட்டுவிட நேரிடும். மாதிரி அளவினைப் பொறுத்து மட்டும் இடர்ப்பாடுகள் ஏற்படுவதில்லை. சில இயல்புகட்கு மற்ற வற்றிலிருந்து மாறுபடும் மாதிரி முறைகள் கையாளப்படுவதாக வும் இருந்திடலாம். அடிக்கடி மாதிரிகளை எடுக்கவேண்டிய சூழ் நிலைகளில், சில இயல்புகளை ஒன்றாக இணைத்து செலவைக் குறைப்பதும், சிறந்த வழி முறைகள் தேவைப்படும் இயல்புகளைத் தனியாகப் பிரித்து கருதுவதும் பயனுடையதாக அமைகிறது. நான்கு விதமாக இயல்புகள் அனைத்தையும் பிரித்துப் பின்வருமாறு கொடுக்கிறோம் :—

	இயல்புகள் அமையும் விதம்.	மாதிரியின் தன்மை
முதல் வகை	அனைத்தும்; வெவ்வேறு பிரிவுகளிலும் தக்க (reasonable) அலைவெண்களுடன் அமையும் நிலை (பரவலாக)	இது அளவில் மாதிரி விகிதத்துடன், பொதுவான ஆய்வு முறை.
இரண்டாம்	வெவ்வேறு பிரிவுகளிலும் குறைந்த அலைவெண்களுடன் அமையும் நிலை.	பெரும் அளவினதான மாதிரி விகிதத்துடன், பொதுவான ஆய்வு.
மூன்றும் வகை	பெரும்பாலான பிரிவுகளில் (Regions) தகுந்த அலைவெண்களுடன் அமைந்து சில இடங்களில் குவிவாகவும் சில இடங்களில் இல்லாமலும் அமையும் நிலை.	வெவ்வேறு பகுதிகளின் முக்கியத்துவத்துக்கு ஏற்ப, கையாளப்படும் படுகை மாதிரி முறை. துணை மாதிரி முறையும் கையாளலாம்.
நான்காம் வகை	சிறுசிறு பகுதிகளில் மட்டும் குவிவான அளவில் பரவல் நிலை	பொது ஆய்வுக்கு பயன்படாது. பரவலின் நிலைக்கு ஏற்ப மாதிரி முறையமைய வேண்டும்.

இப்பிரிவுகள் அனைத்தும், விவசாய ஆய்வுகளின் முடிவுகளின் அடிப்படையில் வகுக்கப்பட்டன. பொதுவான ஆய்வா

னது, கூறுகள் அனைத்தும் ஒழுங்காக பரவலாக அமைந்திருக்கும் ஒருபிரிவில் மேற்கொள்ளப்படும் ஆய்வுமுறை ஆகும். (உதாரணமாக, பின்பு கருதப்படும் எளிய ராண்டம் முறையினைக் குறிப்பிடலாம்.) பொதுவாக, எந்த ஒரு ஆய்விலும் பல்வேறு பிரச்சினைகள் தீவிரமாக அலசி ஆராயப்படுவதால் இத்தகைய ஒரு தெளிவானது தேவைப்படுகிறது. தகுந்த திட்டமிடலானது, ஏற்படும் இழப்பின் அளவினைப் பெரிதும் குறைத்திடக் காணலாம்.

குறியீடுபற்றிய ஒரு விளக்கம்

கணக்கியல் அடிப்படையில் பெறப்படும் சில முடிவுகளிலும், தொடர்களிலும் (results and expressions) பயன்படுத்தப்படும் சில குறியீடுகளைப் பின்வருமாறு பட்டியலிடலாம் :

வரிசை எண் (1)	குறிப்பிடப்படும் இயல்பு (Characteristic) (2)	முழுமைத் தொகுதி குறியீட்டு முறை (Population) (3)	மாதிரியில் குறியீட்டு முறை (Sample) (4)
1.	மொத்த உறுப்புகள் (units)	N	n
2.	கருதியமாறி y-ன் மதிப்புகள்	$Y_1, Y_2 \ldots Y_N$	$y_1, y_2 \ldots y_n$
3.	துணை மாறி x-ன் மதிப்புகள் (Auxiliary)	$X_1, X_2 \ldots X_N$	$x_1, x_2 \ldots x_n$
4.	மாதிரியில் இடம் பெற வாய்ப்புகள்	$\pi_1, \pi_2 \ldots \pi_N$	$\pi_1, \pi_2 \ldots \pi_n$
5.	y-ன் சராசரி மதிப்பு	$\bar{Y}$	$\bar{y}$
6.	விகித சம அளவு (Proportion)	P	p
7.	விகிதம் (Ratio)	R	r
8.	முழுமைத் தொகுதி அளவை & மாதிரி அளவை (parameter & estimate)	θ	t அல்லது $\hat{\theta}$
9.	t என்ற மதிப்பீட்டின் பிறழ்ச்சி (Bias)	$B(t)$	$b(t)$

வரிசை எண் (1)	குறிப்பிடப்படும் இயல்பு (Characteristic) (2)	முழுமைத் தொகுதி குறியீட்டு முறை (Population) (3)	மாதிரியில் குறியீட்டு முறை (Sample) (4)
10.	t மதிப்பீட்டின் மாறுபாடு (Variance)	$V(t$ அல்லது σ^2_t	$v(t)$
11.	t, t' இவற்றின் உடன் மாறுபாடு (Covariance)	$COV\ (t, t')$	அல்லது s_t^2 $cov\ (t, t')$
12	t-ன் தரப்பிழை அளவு (standard error)	σ_t	s_t
13.	t-ன் மாறுபாட்டுக் கெழு (coeficient of variation)	C_t	c_t
14.	t-ன் சராசரிப் பிழைவர்க்கம் (Mean square error)	M_t	m_t
15.	உடன் தொடர்புக் கெழு (Correlation coefficient)	ξ	r அல்லது $\hat{\xi}$
16.	தற்போக்குக் கெழு (regression)	β	$\hat{\beta}$
17.	t' உடன் ஒப்பிட்ட, t-ன் மாதிரி முறைத் திறன் அளவு (Efficiency in sampling)	$E_s\ (t/t')$	$e_s\ (t/t)$
18.	t' உடன் ஒப்பிட்ட, t-ன் செலவு பற்றித் திறன்	$E_c\ (t/t')$	$e_c\ (t/t)$

4. எளிய ராண்டம் மாதிரி முறை அல்லது 'ராண்டம் மாதிரி முறை'

(Simple Random Sampling Method or Random Sampling Method)

[வரையறை — ராண்டம் எண் அட்டவணை — 'ராண்டம் மாதிரி முறை — எடுத்துக் காட்டுகள் — 'ராண்டம்' மாதிரி (ஈடு செய்யப் படாத முறை) மதிப்பீடுகள் — 'ராண்டம்' மாதிரி (ஈடு செய்யப்பட்ட முறை) — சில முடிவுகள் — எடுத்துக் காட்டுகள் மற்றும் பயிற்சிகள் ·வீதம் மற்றும் சதவிதம்' பற்றிய ராண்டம் முறை — எடுத்துக் காட்டுகளும், பயிற்சிகளும் — மாதிரி அளவும், திட்பமும் — தீர் மானங்கள் பற்றிய பிரச்னையில் மாதிரி அளவு.]

வரையறை : நிகழ்தகவின் (probability) அடிப்படையில் அமைந்த சாதாரண மாதிரி முறையினை, அதன் வரையறையின் காரணமாக சாதாரண ராண்டம் மாதிரி முறை என அழைக்கிறோம். இதுவே, முதன்முதலாகக் கண்டு பிடிக்கப்பட்ட மாதிரி முறையும் ஆகும். ஒரு முழுமைத் தொகுதியிலிருந்து அதன் ஒவ்வொரு கூறும் அல்லது அலகும் (unit) தேர்ந்தெடுக்கப்படுவதற்கான சம அளவு வாய்ப்பைப் பெறும் வண்ணம் தெரிவு செய்யப்படும் மாதிரியினை **ராண்டம் மாதிரி** என்று வரையறுக்கிறோம்.

தேர்ந்தெடுக்கப்படும் ஒவ்வொரு நிலையிலும், ஒவ்வொரு கூறும் மாதிரியில் சேர்க்கப்படுவதற்கு சம அளவில் வாய்ப்பு பெற்றிருப்பின், இவ்வாறு எடுக்கப்படும் மாதிரி ராண்டம் மாதிரி யாகவும், இதற்கான வழிமுறையானது **ராண்டம் மாதிரி முறை** எனவும் வரையறுக்கப்படும்.

இத்தகைய மாதிரிக்கு, ஒவ்வொரு கூறும் தெரிந்தெடுக்கப்பட வாய்ப்பினையும், அதுவும் சம அளவினதான வாய்ப்பையும், எடுக் கப்பட்ட கூறு ஈடு செய்யப்படாத நிலையில்(without replacement) ஒன்றினை ஒன்று சார்ந்த வாய்ப்புகள் பெற்ற கூறுகளைத் தேர்ந் தெடுப்பதனையும், கருத்தில் கொள்ளல் வேண்டும். இது ராண்டம் மாதிரியினை அளிக்கும். மாறாக எடுக்கப்பட்ட கூறு ஈடுசெய்யப்

பட்ட அளவில் (with replacement) பெறப்படும் மாதிரியானது எளிய ராண்டம் மாதிரி (simple random sample) எனப்படும். எடுக்கப்படும் ராண்டம் மாதிரிகளெல்லாமே, எளிய ராண்டம் மாதிரியாக அமைதல் இயலாது. இங்கொன்று அங்கொன்ருக எடுக்கப்படும் மாதிரியும் ராண்டம் மாதிரி ஆகாது ஒரு குறிப்பிட்ட அடிப்படையில் மாதிரித் தெரிவு அமையவேண்டும்.

ராண்டம் மாதிரியானது. பின் வரும் வழிகளில் தேர்ந்தெடுக்கப்படலாம் :

(1) 'லாட்டரி' முறை. (Lottery Method)

(2) 'ராண்டம்' எண் பட்டியலை (அட்டவணை) பயன்படுத்துதல் (Use of random numbers table).

முதலாம் முறைப்படி, N அளவுள்ள ஒரு முழுமைத் தொகுதியிலிருந்து, n அளவுள்ள ஒரு மாதிரியினை தேர்ந்தெடுப்பதாகக் கருதுவோம். முதலில், '1' முதல் 'N' வரையிலான எண்களை முழுமைத் தொகுதியின் N கூறுகளுக்கும் அளிக்க வேண்டும். பிறகு இந்த எல்லா எண்களையும் வெவ்வேறு (ஒரே வடிவுடைய) காகிதத் துண்டுகளில் (slip) எழுதுதல் வேண்டும். இத்தகைய துண்டுகள் அல்லது அட்டைகள் (card) ஒரே மாதிரி இருக்க வேண்டியது அவசியம். பிறகு இந்த அட்டைகளை ஒரு பெட்டியில் இட்டு நன்கு குலுக்கவேண்டும். நன்கு கலக்கப்பட்டு விடுவதால், எந்தக் குறிப்பிட்ட சீட்டுக்கும் அல்லது அட்டைக்கும் அதிக அளவில் தேர்வு வாய்ப்பு ஏற்படுவதில்லை. அதன் பின்னர், சீட்டுக்கள் ஒன்றன்பின் ஒன்ருய்த் தேர்ந்தெடுக்கப்படுகின்றன. ஒவ்வொரு முறையும் சீட்டுகள் நன்கு கலக்கப்படுதல் மிகவும் இன்றியமையாதது. முதலில் தேர்வுபெறும் 'n' சீட்டுகள் ஈமக்குத் தேவைப்பட்ட மாதிரியாக அமையும். முழுமைத் தொகுதியின் அளவு பெரிதாக இருப்பின் இம்முறையினைப் பின்பற்றுவது மிகக் கடினம்.

ராண்டம் எண் அட்டவணை (Random Numbers Table)

ராண்டம் மாதிரியினைத் தேர்ந்தெடுக்க இது மற்ருெரு முறை. ராண்டம் எண் பட்டியல் அல்லது அட்டவணை மிக விரிவான அளவில் பயன்படுத்தப்படுகிறது. இத்தகைய எண்கள் பட்டியல் பலவகைப்படும் :

(i) கெண்டால், ஸ்மித் ராண்டம் எண் பட்டியல். (Kendall & Smith Random No. Table)

(ii) ‘ டிப்பெட் ’—ன் ராண்டம் எண் பட்டியல்.
(Tippett's Random Number Table)

(iii) ‘ ஃபிஷர் — ஏட்ஸ் ’ ராண்டம் எண் பட்டியல்.
(Fisher & Yates Random Number Table)

(iv) பத்துலட்சம் எண் கொண்ட ‘ ராண்ட் கார்ப்பொரேஷன் ’ ராண்டம் எண் பட்டியல்
(Rand Corporation Random Nos. Table)

‘ டிப்பெட் ’ பட்டியலில் மொத்தம் 41,600 இலக்கங்கள் உள்ளன. இவற்றை 10,400 நான்கு இலக்கமுள்ள எண்களாகவும் கருதலாம்.

ராண்டம் எண் அட்டவணையினைப் பயன்படுத்தும் முறை

N உறுப்புகள் கொண்ட ஒரு முழுமைத் தொகுதியிலிருந்து *n* உறுப்புகள் உள்ள ஒரு மாதிரியினைத் தேர்ந்தெடுப்பதாகக் கொள்வோம். இதற்கு அடிப்படையான ‘ சட்ட ’ வரம்பினை (frame) நிர்ணயம் செய்த பின்னர், 1 முதல் *N* வரையிலான எண்களை முழுமைத் தொகுதிக் கூறுகளுக்கு வழங்கவேண்டும் உதாரணமாக, $N=500$ எனில். வழங்கப்பெறும் எண்கள் 001-முதல் 500 வரையில் ஆகும். ஏதாவது ஒரு நிரலினை (column) ராண்டம் எண் பட்டியலிலிருந்து கருதி தொடர்ச்சியாக, குறிப்பிட்ட இலக்கங்கள் கொண்ட எண்களைத் தெரிவு செய்ய வேண்டும். மேற்கண்ட உதாரணத்தில், மூன்று இலக்கமுள்ள எண்களை கருதுதல் அவசியம். மேலிருந்து கீழான முறையில் (vertically) அல்லது இடமிருந்து வலமான முறையில் (horizontally) எண்களைத் தேர்ந்தெடுக்கலாம். ஏதாவது எண் 500 க்கு மேலாக இருப்பின், (உதாரணம்), அதனை நீக்கிவிடவேண்டும். ஒவ்வொரு எண்ணும் ஒரு முறைதான் கணக்கிடப்பட வேண்டும். முழுமைத் தொகுதி அளவு *N*-ஐ விட பெரிதாக எண்கள் அமைந்திருப்பின், அதிக அளவில் எண்களை நீக்கவேண்டியிருக்கும். எனவே *N* என்ற மதிப்பின் பெருக்கங்களை (multiples) எடுக்கப்பட்ட எண்களிலிருந்து கழித்து மீதி மதிப்புக்களைக் கருதலாம்.

எடுத்துக் காட்டு (1):

400 உறுப்புகள் கொண்ட முழுமைத் தொகுதியினின்று, 10 உறுப்புகளுள்ள ஒரு மாதிரியினைத் தேர்ந்தெடுத்தல்:

தேவையான மாதிரிக்கூறுகளை நிர்ணயித்து 1 முதல் 400 வரையில் வரிசையாக எண்களை, முழுமைத் தொகுதிக் கூறு களுக்கு வழங்க வேண்டும். ராண்டம் எண் பட்டியலிலிருந்து 10 மூன்று இலக்க எண்களைத் தேர்ந்தெடுத்து எழுதுதல் வேண்டும். பின்வரும் எண்களைக் கருதலாம் :

537	634	645	348
633	982	850	
353	026	585	

சில எண்கள் 400 க்கு மேல் உள்ளதால் 400-ன் பெருக் கங்களை (multiples) இந்த எண்களினின்று கழிக்க; பெறப்படும் எண்கள் தேவைப்படும் மாதிரிக்கூறினை அமைக்கும்:

137	234	245	348
233	182	050	
353	026	185	

மேலே குறிப்பிட்ட எண்களை வரிசை எண்களாகப் பெற்ற கூறுகள் (units) மாதிரியினை அளிக்கும். இந்த மாதிரியானது ராண்டம் மாதிரி (random sample) ஆகும்.

குறிப்பு : இவ்வாறு தேர்ந்தெடுக்கப்படும் எண்களுக்கு உரிய நபர்களுடன் (அல்லது மாதிரி அலகுகளுடன்) தொடர்பு கொள்ள இயலாவிடில் அவருக்குப் பதிலாக வேறொருவரைக் கருது வதற்கு ஏற்ற வகையில் முன் கூட்டியே அதிகப்படியான 'ராண்டம்' எண்களைத் தெரிவு செய்து கொள்ளுதல் நன்று. அதாவது 10 எண்களுக்கு பதில், 15 எண்களைத் தேர்ந்தெடுத்துக் கொள்ளலாம். இது நடைமுறையில் உதவத்தக்க குறிப்பாகும்.

எடுத்துக்காட்டு (2) :

ஒரு தொகுதியில் உள்ள 15 தெருக்களில் உள்ள வீடுகளின் எண்ணிக்கைகள் வருமாறு :

15 25 18 20 35 27 14 16 12 16
32 28 10 14 13.

மேற்கண்ட 295 வீடுகளிலிருந்து; 10 வீடுகளைக் கொண்ட மாதிரியினைத் தேர்ந்தெடுத்தல் :

வீடுகளின் தொடர்ச்சியான கீழீட்டுக் கூட்டு மதிப்புக்கள் (sub-totals) பின்வருமாறு அமையும் :

15 40 58 78 113 140 154 170 182 198

230 258 268 282 295.

ராண்டம் எண் பட்டியலிலிருந்து 10 எண்களைத் தெரிவு செய்து (மூன்று இலக்கமுள்ளவை) எழுதலாம்.

613 990 067 655 019 715 482 011 515 614.

நிறைய எண்கள் 295-ஐ விட அதிகமாய் இருப்பதால், அவற்றிலிருந்து 300-ன் பெருக்கங்களைக் கழிக்க.

013 090 067 055 019 115 182 011 215 014 என்று எண்கள் அமைந்திடும். ஏதாவது எண் இரண்டாம் முறை திரும்ப வர நேரிடின் அதை நீக்கி அதற்கு அடுத்த எண்ணைத் தேர்ந்தெடுக்க வேண்டும்.

தெரு எண் வரிசைப்படி அட்டவணையினை அமைத்து மாதிரிக் கூறுகளை விவரமாகக் குறிக்கலாம்.

முதல் எண் 013 என்பது முதலாவது தெருவிலேயே அமைகிறது. ஏனெனில் 1 முதல் 15 எண்கள் வரையில் உள்ள 15 வீடுகள் முதல் தெருவில் அடங்கும். 16 முதல் 40 வரை இரண்டாவது தெருவில் அடங்கும். இதேபோல் ஏனைய ராண்டம் எண்களையும் பிரிவு படுத்த வேண்டும். இதன் மூலம் தெரிவு செய்யப்படும் மாதிரியானது பின்வருமாறு அட்டவணைப் படுத்தப்படலாம் : (பக்கம் 73)

எடுத்துக்காட்டு: 3: 10 எண்கள் கொண்ட பிரிவுகளாக (Classes) அமைக்கப்பட்ட அலைவெண் பட்டியிலினை (frequency distribution) ராண்டம் எண் பட்டியலிலிருந்து எடுக்கப்பட்ட 150 இரு இலக்க எண்களின் பெருக்கற் பலன்களுக்கு (Product) வரையறை செய்து, கண்டறியப்பட்ட மதிப்புகளையும் எதிர் நோக்கும் மதிப்புகளையும் (observed and theoretical values) ஒப்பிட்டு ஆய்க. (B.Sc செ. 1963)

தெரு எண் :	1	2	3	4	5	6	7
வீடுகளின் எண்ணிக்கை	1—15	16—40	41—58	59—78	79—113	114—140	141—154
மாதிரி	(11, 13, 14) 11வது, 13வது 14வதுவீடுகள்	(19) 4-வது வீடு	(55) 15-வது வீடு	(67) 9-வது வீடு	(90) 12-வது வீடு	(115) 2-வது வீடு	

தெரு எண்	8	9	10	11	12	13	14	15
வீடுகளின் எண்ணிக்கை	155—170	171—182	183—198	199—230	231—258	259 -268	269—282	283—295
மாதிரி		(182) 12-வது வீடு		(215) 17-வது வீடு				

அந்தந்தப் பிரிவுகளில் அமையும் வீடுகளின் இலக்கத்திற்கு ஏற்ற வகையில், மாதிரித் தேர்வில் முறையே, 1-வது, 2-வது, 3-வது, 4-வது, 5-வது, 6-வது, 9-வது, 11-வது தெருக்களில், அந்தந்தப் பிரிவில் காட்டப் பட்டபடி வீடுகளைக் கருத வேண்டும். இது ராண்டம் முறையில் நன்கு அமைந்திருப்பதைக் கண்கூடாகக் காணலாம். எல்லாத் தெருக்களும் கலந்த ஒரு அமைப்பில் கூறு எடுக்கப்படுகிறது. ஒரே தன்மையான (homogeneous) முழுமைக் கூறு அமைந்திருத்தல் இன்றியமையாதது.

ராண்டம் எண் அட்டவணையிலிருந்து, 150 இரு இலக்க எண்களைத் தெரிவு செய்வோம் :

61	19	35	18	99	10	20	80	58	00
99	13	43	24	75	93	28	28	75	33
06	28	80	54	21	64	57	07	30	81
65	22	20	38	44	74	79	44	72	14
01	24	92	08	10	73	02	64	94	76
71	66	13	32	11	01	74	16	22	14
48	26	82	00	52	98	36	41	35	81
01	23	40	72	40	03	97	75	58	75
51	60	44	51	01	02	46	75	31	81
61	43	71	91	02	79	22	19	52	39
57	70	56	50	06	05	28	58	20	72
37	98	06	22	63	88	65	53	74	59
60	89	31	10	22	80	17	19	54	20
45	79	37	54	20	91	18	11	44	82
51	03	58	48	89	32	82	94	64	23

('டிப்பட்' ராண்டம் எண் பட்டியலில் இருந்து எடுக்கப்பட்ட எண்களாகும்.)

மேற்கண்ட 150எண்களில் ஒவ்வொன்றிலும் உள்ள இலக்கங்களின் பெருக்கற் பலனைக் கணக்கிட்டுக் கிடைக்கும் எண்களை வேறொரு பட்டியலில் அமைத்திடுவோம். உதாரணமாக; '61' என்னும் எண்ணிக்கை உள்ள இலக்கங்களின் (Digits) பெருக்கற் பலன் $6 \times 1 = 6$ என்பதாகும். இதே போல் எல்லா எண்களுக்கும் கணக்கிடல்வேண்டும். மொத்தம் 150 பெருக்கற் பலன்களைக் கொண்டு அமைத்த பட்டியல் வருமாறு:—

6	9	15	8	81	0	0	0	40	0
81	3	9	8	35	27	16	16	35	9
0	16	0	20	2	24	35	0	0	8
30	4	0	24	16	28	63	16	14	4
0	8	18	0	0	21	0	24	36	42
7	36	3	6	1	0	28	6	4	4
32	12	16	0	10	72	18	4	15	8
0	6	0	14	0	0	63	35	40	35
5	0	16	5	0	0	24	35	3	8
6	12	7	9	0	63	4	9	10	27
35	0	30	0	0	0	16	40	0	14
21	72	0	4	18	64	30	15	28	45
0	72	3	0	4	0	7	9	20	0
20	63	21	20	0	9	8	1	16	16
5	0	40	32	72	6	16	36	24	6

இரண்டு ராண்டம் இலக்கங்களின் (random digit) பெருக்கற் பலனின் மிகக் குறைந்த மதிப்பு (minimum value) ஆனது '0' ஆகவும் மிக உயர்ந்த மதிப்பு (maximum value) ஆனது 9 9=81 எனவும் அமைகின்றன. எனவே, பிரித்து (classified) வகை செய்யப்பட்ட அலைவெண் பட்டியலானது 10 என்ற பிரிவு இடைவெளி (class interval)-யினைக்கொண்டு, மேற்கண்ட 150 மதிப்புகளையும் கொண்டதாக அமைகிறது. அப்பட்டியலாவது :

பிரிவு எல்லை. (1)	கண்டறிந்த அலைவு. (2)	மொத்தம் சார்ந்த அலைவு வீதம் கண்ட றிந்தவை) (3)	எதிர் நோக்கிய அலைவு வீதம் (விதிப்படி) (4)
0—10	77	0·5133	0 42
10—20	24	0·1600	0·17
20—30	17	0·1133	0·13
30—40	15	0·1000	0·09
40—50	6	0·0400	0·09
50—60	0	0·0000	0.04
60—70	5	0 0333	0·03
70—80	4	0·0267	0·02
80—90	2	0 0134	0·01
மொத்தம்	150	1·0000	1·00

அலைவு வீதம் (relative frequency) ஒவ்வொரு பிரிவிலும் (class interval) உள்ள அலைவு (frequency) எண்ணிற்கும் மொத்த அலைவு (frequency total) மதிப்பிற்கும் இடையே உள்ள விகித அளவாக கணக்கிடப்படும். எதிர்நோக்கும் அல்லது எதிர் பார்க்கும் அலைவு வீதம் (Expected or theoretical relative frequency) வாய்ப்பு விதி நியதியின் (Probability Basis) அடிப் படையில் கணிக்கப்படும். 0—10 எனும் பிரிவில் '0' முதல் '9' வரை அமையும் பெருக்கற் பலன்களைத் தரும் (product) இரு இலக்க எண்கள்

00	10	20	41	90
01	11	21	42	91
02	12	22	50	
03	13	23	51	
04	14	24	60	
05	15	30	61	
06	16	31	70	
07	17	32	71	
08	18	33	80	
09	19	40	81	

என்ற '42' இரு இலக்க எண்களாகக் காணலாம். இந் நிகழ்ச்சி அமைய மொத்த சாத்தியக் கூறுகள் (total possibilities) '100' ஆகும். எனவே, வாய்வு விதிப்படி 0—10 என்ற பிரிவுக்கேற்ற எதிர்பார்க்கும் (theoretical) அலைவீதம் $\frac{42}{100}$ =0·42 ஆகும், இதேபோல, மீதி எட்டுப் பிரிவுகட்கும் ஏற்ற எதிர்நோக்கி அலைவு வீதங்கள் கணிக்கப்படுகின்றன. பத்தி (4)-ல் இவை குறிக்கப்பட்டுள்ளன. (3)-ம், (4)-ம் (column) ஒப்பிட்டு நோக்கத்தக்கவை. அதன்மூலம் நடைமுறையானது (practice), விதிமுறை (theory)-யிலிருந்து எந்த அளவிற்கு வேறுபட்டுள்ளது என்பதை புள்ளியியல் அடிப்படையில் முடிவு செய்ய இயல்கிறது. இரண்டு பத்தி [(3) & (4)]-களின் கூட்டுத் தொகைகளும் (totals) சமமாக அமைவது குறிப்பிடத்தக்கது.

இதேபோல, குறிப்பிட்ட முறைகள்மூலம் அளந்தறிய இயலாத, பண்புகள் (attributes) அடிப்படையில் அமைந்த முழுமைத் தொகுதிகளிலிருந்தும் மாதிரிக் கூறுகள் தெரிவுசெய்ய இயலும்.

எடுத்துக்காட்டு (4)

பின்வரும் பட்டியலானது, (அட்டவணை-Table) கருதப்பட்ட 750 நபர்களுள், மலேரியா வியாதித் தடுப்பு ஊசி போடப்பட்டதற்கும், நோயினுல் பாதிக்கப்படாமைக்கும் (exemption) உள்ள தொடர்பு நிலையினை (Association) காட்டுகிறது. இப்பட்டியலி னின்று (Table), 40 நபர்கள் கொண்ட மாதிரியினைத் தெரிவு செய்க.

	நோய் மூலம் பாதிக்கப் படாதோர்	பாதிக்கப் பட்டோர்	மொத்தம்
ஊசி போடப்பட்டோர்	266 (i)	5 (ii)	270
போடப்படாதோர்	430 (iii)	50 (iv)	480
மொத்தம்	695	55	750

'ராண்டம்' எண் பட்டியலின்மூலம், 750 எனும் எண்ணுக்குக் கீழாக, அல்லது சமமாக, (less than or equal) மூன்று

இலக்கமுள்ள, 40 ராண்டம் எண்களைத் தேர்வு செய்வது முதல்படி (First step),

எடுக்கப்பட்ட எண்கள் :

028	158	623	421
252	501	401	387
613	251	648	157
747	719	285	321
381	383	723	267
615	681	343	290
596	584	742	131
626	566	622	112
303	214	519	570
249	005	732	413

அடுத்த கட்டம் (Second step), மேற்கண்ட எண்களைப், பின்வருமாறு அட்டவணைப்படுத்தவேண்டும்:

உட்பிரிவு (cell)	பிரிவு எல்லை (class limits)	கணிப்புக் குறி (tally mark)	உட்பிரிவு அலைவெண் (call frequency)
(I)	1 —265	卌 卌	10
(ii)	266—270	|	1
(iii)	271—700	卌 卌 卌 卌 ||||	24
(vi)	701—750	卌	5
			40

மேற்கண்ட 40 ராண்டம் எண்களை, கீழ்வரும் தேர்வுப் பட்டியலில் (Contingency) வகைப்படுத்துவது, மூன்றாவது நிலை (Third Stage),

	பாதிக்கப் படாதோர்	பாதிக்கப் பட்டோர்	மொத்தம்
ஊசி போடப்பட்டோர்	10	1	11
ஊசி போடப்படாதோர்	24	5	29
மொத்தம்	34	6	40

வெவ்வேறு உட்பிரிவுகளுக்கான, கண்டறிந்த (observed) நிகழ்தகவுகளும் (Probabilities), எதிர்பார்க்கும் (expected) நிகழ்தகவுகளும் வருமாறு:

உட்பிரிவு	மாதிரி (Sample)	முழுமைத் தொகுதி (population)
(i)	0· 250	0· 353
(ii)	0 025	0· 007
(ii)	0 600	0· 573
(iv)	0· 125	0· 073

முழுமைத் தொகுதி மதிப்புகள் எதிர் நோக்கப்படுபவை (Theoretical) மாதிரி அளவீடுகள் உண்மையில் கண்டறியப் படுபவை. இரண்டிற்குமுள்ள வேறுபாடு நன்கு தெளிவாகும்.

இனி, சில புள்ளியியல் பரவல்களை (நிகழ்தகவின் அடிப் படையில்) கருதி, அவற்றிற்கு ஏற்ப அமைகின்ற முழுமைத் தொகுதிகளிலிருந்து ராண்டம் மாதிரிகளைத் தெரிவு செய்யும் வழிமுறையினைப் பின்வரும் மூன்று எடுத்துக்காட்டுகள்மூலம் விளக்கலாம். இவற்றில் கருதப்பட்ட பரவல்கள் (distributions) முறையே

(i) ஈருறுப்பு பரவல் (Binomial distaibution)

(ii) பாய்ஸான் பரவல் (Poisson distribution)

(iii) இயல்நிலைப் பரவல் (Normal distribution)

என்பனவாகும்.

எடுத்துக்காட்டு (i) :

ராண்டம் எண் பட்டியலிலிருந்து, ஐந்து இலக்கங்கள் தொடர்ந்து வரும்படியாக, 30 எண்களைத் தெரிவு செய்க. 0, 1, 2, என்ற எண்களை (இலக்கங்களை) வெற்றி எனவும், மீதமுள்ள இலக்கங்களைத் தோல்வி எனவும் கொள்க. இவ்வாறு கணக்கிடப் பட்ட வெற்றிகளின் (நிகழ்வு) எண்ணிக்கைக்கு உரிய பரவலினை தெரிந்தெடுத்த 30 எண்களுக்கும் அமைக்க. இப் பரவலின் சராசரி, தரவிலக்கம் ஆகியவற்றைக் காண்க. அம் மதிப்புக்களை, நியதிப்படி பெறப்பட்ட மதிப்புக்களுடன் ஒப்பிடுக.

ஐந்து இலக்க ராண்டம் எண்கள் முப்பதை ($n=30$), ராண்டம் எண் அட்டவணையினின்று தெரிவு செய்க. அவற்றைப் பின்வருமாறு கொடுக்கிறோம் :

61196	32179	98520
15474	69234	11805
94557	19565	83452
42481	45155	88685
23523	94864	99594
04493	98086	74029
00549	33185	54178
35963	80951	11664
59808	79752	48324
46058	18633	69074

ஒவ்வொரு எண்ணும் ஐந்து இலக்கமுடையதாக உள்ளதைக் காண்க. ஒவ்வொரு ராண்டம் எண்ணிலும், 0, 1, 2 என்ற இலக்கங்கள் வரின், அதனை 'வெற்றி' (அல்லது நிகழ்வு) என்றும் (Success), மீதி ஏழு எண்களில் (அதாவது 3, 4, 5, 6, 7, 8, 9 என்பவை) ஏதேனும் வரின், அதனை 'தோல்வி' (அல்லது நிகழாதது) எனவும் கருதுக. மேற்கண்ட எண்களில்

ஏதேனும் ஒன்றில் ஒரே இலக்க மதிப்புத் திரும்பவும் வரின், இரண்டு நிகழ்வுகள் அல்லது வெற்றிகள் அமைவதாகக் கொள்ள வேண்டும். உதாரணமாக, '18533' என்ற எண்ணில், '1' என்ற இலக்கம் ஒருமுறை வருவதால், ஒரே ஒரு வெற்றி அமைவதாக கொள்கிறோம். மற்ற இலக்கங்கள், '2'க்கு மேல் இருப்பதால், அவை தோல்விகளாகக் கருதப்படுகின்றன, ($x=1$) என்பதில் ஒரு குறியீடு அமைக்கப்படுகிறது (/ என்ற குறியீடு). ஒவ்வொரு எண்ணும் ஐந்து இலக்கங்கள் உள்ளதாய் அமைவதால், அதிகபட்சமாக ஐந்து வெற்றிகளை (அல்லது நிகழ்வுகளை) பெறலாம். இதன் குறைந்தபட்ச மதிப்பு பூச்சியமாகும். ∴ x-என்ற வாய்ப்பு மாறியானது, பெறக்கூடிய வெற்றிகளை (நிகழ்வுகளைக்) குறிப்பதாகக் கொண்டு, $x = 0, 1, 2, 3, 4, 5$ மதிப்புகள் வரக் காணலாம்.

கருதப்பட்ட 30 ராண்டம் எண்களைக் கொண்டு, மொத்தம் வரக்கூடிய வெற்றிகளை, ஒரு ஈருறுப்புப் பரவலினின்று (binomial distribution) எடுக்கப்படும் மாதிரியாக (சம வாய்ப்பு என்ற அடிப்படையில்) பின்வருமாறு அட்டவணைப்படுத்தலாம்:

பெறக்கூடிய வெற்றிகள் (x)	குறியீடு	மொத்த எண்கள் (f)
0	𝍸	5
1	𝍸 𝍸 𝍸 |	16
2	𝍸 |||	8
3	/	1
4		0
5		0
		மொத்தம் 30

இந்த மாதிரியின் சராசரி மற்றும் மாறுபாட்டு அளவு ஆகியவை

$$\bar{x} = \frac{\Sigma fx}{\Sigma f}$$

$$= \frac{(5 \times 0) + (16 \times 1) + (8 \times 2) + (1 \times 3)}{30}$$

$$= \frac{35}{30} = 1{\cdot}17$$

$$S^2 = \frac{\Sigma f x^2}{\Sigma f} - (\bar{x})^2$$

$$= \frac{16 + 32 + 9}{30} - (1.17)^2$$

$\therefore S^2 =$ 5. 53 என்று பெறப்படுகின்றன.

முழுமைத் தொகுதியில் (ஈருறுப்புப் பரவல் அடிப்படையில்), பின்வருமாறு கருதுக:

மொத்தம் பத்து இலக்கங்கள் அமைய வாய்ப்பு உள்ளது. அவை 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 என்பவை. அவற்றுள் மூன்று குறிப்பிட்ட இலக்கங்களான 0, 1, 2 ஆகியவற்றை 'வெற்றி' அல்லது நிகழ்வு (success) எனக் கொள்வதால், இதற்கேற்ற வாய்ப்பு நிலை

$p = \frac{3}{10} = 0 \cdot 3.$ ஆக அமைகிறது. எனவே,

$q = \frac{7}{10} = 0 \cdot 7$ என வருகிறது.

இவற்றைப் பயன்படுத்தி, முழுமைத் தொகுதியின் சராசரி, மற்றும் மாறுபாடு ஆகியவற்றை

$np = 5 \times 0 \cdot 3 \qquad (n = 5)$

$= 1 \cdot 5$ என்றும்,

$npq = (1 \cdot 5).(0 \cdot 7) = 1 \cdot 05$

என்றும் பெறப்படும்.

எனவே, எதிர்பார்க்கும் அலைவெண் மதிப்புகளை (expected frequencies), ஈருறுப்புப் பரவலான

$$N.f(x) = 30. \binom{5}{x}. (0 \cdot 3)^x . (0 \cdot 7)^{5-x}$$

என்ற வரையறைப்படி $x = 0, 1, 2, 3, 4, 5$ என்ற மதிப்புகட்குப் பெறலாம்.

பயிற்சி: இவற்றைக் கணக்கிட்டு, பட்டியல் அமைத்திடுக.

எடுத்துக்காட்டு: (ii) 1000 கூறுகள் கருதப்பட்ட ஒரு பாய்ஸான் பரவலின் (Poissin distribution) ஒழுங்கு விதிப்படி எதிர்பார்க்கும் அலைவெண் மதிப்புகளைக் கணக்கிடுதல். சராசரியுமை 2 என்று கொண்டு, 10 உறுப்புகள் கொண்ட ஒரு சமவாய்ப்பு மாதிரியினைத் தெரிவு செய்து அதன் சராசரி, தரவிலக்கம் ஆகியவற்றின் மதிப்பீடுகளைப் பெறுக. அம் மதிப்புக்களை முழுமைத் தொகுதி மதிப்புக்களுடன் ஒப்பிடுக.

1000 என்ற எண்ணிக்கை அமைந்த ($N = 1000$) பாய்ஸான் பரவலின் எதிர்பார்க்கும் அலைவு எண்மதிப்புக்களை, $m = 2$ என்று கொண்டு கணக்கிடுகிறோம்.

பாய்ஸான் விதியானது

$$f(x) = \frac{e^{-m}.m^x}{x!} \quad x = 0, 1, 2 \ldots\ldots$$

என்பதாகும். இதில் $m = 2$ என்று பிரதியிட,

$$f(x) = \frac{e.^{-2}\ 2^x}{x!} \text{ என்ற விதியினில்,}$$

$x = 0, 1, 2 \ldots\ldots 9$ ஆகிய மதிப்புகட்கு எதிர்பார்க்கும் அலைவெண் மதிப்பீடுகளைப் பெறலாம்.

x	$f(x)$	எதிர்பார்க்கும் அலைவெண்கள் $N.\ f(x)$
0	0·135	135
1	0·271	271
2	0·271	271
3	0·181	180
4	0·090	91
5	0·036	36
6	0·012	12
7	0·003	3
8	0·001	1
9	0·0002	0
	$N;$	1000

இந்தப் பரவலின் சராசரி, மற்றும் மாறுபாட்டு அளவு ஆகியவற்றின் மதிப்புக்கள்

$$\mu = 2 \quad ; \quad \sigma^2 = 1·9$$

என்பவையாகப் பெறப்படும்.

மேற்குறிப்பிடப்பட்ட $m = 2$ என்ற நிலையில் அமைந்த பாய்ஸான் பரவலிருந்து ஒரு சம வாய்ப்பு மாதிரியினை, $n = 10$ என்ற அளவில் தெரிவு செய்கிறோம். இதற்கு, ராண்டம் எண்

களின் பட்டியலைக் கருதி. 1000 என்ற எண்ணுக்குக் குறைவாக (1000 உள்பட) அமையும் பத்து எண்களைத் தெரிவு செய்க.

தெரிவு செய்யப்பட்ட எண்கள்

263, 392, 845, 708, 185, 728, 296, 845, 686, 306. என்பதாகக் கொள்வோம்.

இவற்றை அலைவெண்களாகக் கருதி, தொடர்ந்து கூட்டப் பெற்று வரும் அலைவெண்கள் கொண்ட (cumulative frequency) பட்டியலினை உருவாக்குகிறோம். மாதிரியினைக் கருதிய அட்டவணையாவது, அலைவெண்களைக் கொண்டு பின்வருமாறு அமைகிறது:

x	அலைவு எண் (Frequency)	தொடர் கூட்டுஅலைவு (Cumulative freq.)	'/' குறியீடு	மாதிரி அலைவு (Sample freq.)
0	135	135		0
1	271	406	𝍸	5
2	271	677		0
3	181	858	𝍸	5
4	90	948		0
5	36	984		0
6	12	996		0
7	3	999		0
8	1	1000		0
	1000			10

[கருதப்படும் ராண்டம் எண், 1—135 என்ற பிரிவில் வரின் $x=0$-க்கு நேராகவும், 136—406 என்ற பிரிவில் வரின், $x=1$-க்கு நேராகவும், இதேபோல் எல்லா எண்களையும் தொடர்ந்து கருதி, குறியீடு செய்யப்படுகிறது. இவற்றின் மூலம், மாதிரி அலைவுகள் x—ன் வெவ்வேறு மதிப்புகட்குப் பெறப்பட்டு, மாதிரி அலைவுப் பட்டியல் (sample frequency table) பூர்த்தி செய்யப்படுகிறது.]

இந்த மாதிரியின் சராசரி, மற்றும் மாறுபாடு இவற்றின் மதிப்புக்கள் :

சராசரி : $\bar{x} = \frac{20}{10} = 2$

மாறுபாடு : $s^2 = \frac{50}{10} - (4\cdot0) = 5-4 =$

s^2 மதிப்பு, σ^2 மதிப்பினின்று வேறுபட்டிருப்பது காணத் தக்கது. மாதிரி அளவினை அதிகப்படுத்துகையில், இவ்வேறு பாடு குறையக் காணலாம்.

இனி, மாதிரி ஆய்வுகளில் பெரும்பாலும் பயன்படும் இயல் நிலைப் பரவலின் (normal distribution) அடிப்படையில் சம வாய்ப்பு மாதிரி தெரிவு செய்யப்படுவதைக் காண்போம்:

எடுத்துக்காட்டு (iii).

சராசரி $\mu=50$ எனவும், தரவிலக்கம் $\sigma=10$ எனவும் அமைந்த, இயல்நிலைப் பரவலினை (normal distribution) அடிப்படையாகக் கொண்ட முழுமைத் தொகுதியிலிருந்து $n=25$ என்ற மாதிரியினைத் தெரிவு செய்து, அதன் சராசரி மற்றும் தரவிலக்க மதிப்புகளைக் காண்க :

முதல் நிலை : ராண்டம் எண் பட்டியலைப் பயன்படுத்தி, நான்கு இலக்கங்கள் கொண்ட 25 எண்களைத் தெரிவு செய்கிறோம். இவ் வெண்களை தசம பின்னங்களாக (decimals) கருதுகிறோம். அவை பின்வரும் பட்டியலில் அமைகின்றன.

0·6132	0·5799	0·6606
0·9900	0·3775	0·2663
0·0672	0·6021	0·2322
0·6551	0·4544	0·6020
0·0191	0·5110	0·4389
0·7150	0·1911	
0·4822	0·1352	
0·0110	0·2840	
0·5154	0·2201	
0·6148	0·2402	

இரண்டாம் நிலை : மேற்கண்ட ஒவ்வொரு மதிப்பும், இயல் நிலை வளைவில் (normal curve) $(-\infty, t)$ என்ற பிரிவு இடைவெளியில் (interval) காணத்தக்க நிகழ்தகவின் (probability) அளவாகக் கருதப்பட்டு.

$$\int_{-\infty}^{t} \phi(t)\, dt = \alpha$$

என்று குறிப்பிடப்படுவதாகும்.

$$\text{இங்கு } \varphi(t) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2} \text{ ஆகும்.}$$

α என்னும் கெழுவானது பட்டியலில் அமையும் வெவ்வே, வாய்ப்பு அளவுகளைக் குறிக்கிறது. ϕ (t) என்ற சார்பானது இயல்நிலைப் பரவலில் தரப்படுத்தப்பட்ட மாறியான (standard-normal variate), t=யின் நிகழ்தகவுச் சார்பாக (probability function) அறியப்படும்.

மூன்றும்நிலை : இயல்நிலைப் பரவலில் வாய்ப்பு (நிகழ் தகவு) மதிப்புகளைக் காட்டும் அட்டவணையினைப் பயன்படுத்தி ஒவ்வொரு α-விற்கும் ஏற்ற t-யின் மதிப்பினை அறிகிறோம். எனவே முதற்பட்டியலில் உள்ள வாய்ப்புகட்கு ஏற்ற t-மாறியின் மதிப்புகள் பின்வருமாறு பட்டியலிடப்படும் :

'**t**' **மதிப்பீடுகள் : (இயல்நிலைப் பரவல் பரப்பகளைக்கருதி)**

0.288	0.202	0.414
2.330	—0.312	—0.969
—1.497	0.259	—0.731
0.399	—0.115	0.258
—2.071	0.028	—0.156
0.392	—1.334	
—0.045	—1.103	
—3·10	—0·571	
—0.038	—0.772	
0.392	—1.080	

நான்காம் நிலை : பின்வரும் வரையறைமூலம்,

$$= t \quad \frac{x-\mu}{\sigma}$$

$$\left.\begin{array}{lcl} \text{அல்லது } x & = & \mu + t.\sigma \\ \text{அல்லது } x & = & 50 + (t \times 10) \\ \therefore \quad x & = & 50 + 10.t \end{array}\right\} \quad \begin{array}{l} \mu = 50 \\ \sigma = 10 \end{array}$$

என்றவாறு, ஒவ்வொரு t மதிப்பிற்கும் ஏற்ற x மதிப்பினைக் காணலாம்.

எடுத்துக்காட்டாக, $t = 0{\cdot}288$ எனில், $x = 50 + (10 \times .288)$
$= 52.88$ ஆகும்.

∴ x மதிப்புக்கள் :

52.88	53.92	52.02	36.66	54.14
73.30	49.55	46.88	38.97	40.31
35.03	19.00	52.59	44 29	42.69
53.99	50.38	48.85	41.28	52.58
29.29	46.08	50.28	39.20	48.44

இறுதி நிலையில், அலைவெண் அட்டவணை (frequency table) அமைக்கப்படுகிறது. பிரிவு இடைவெளி '8' எனக் கொண்டு பெறப்படுகிறது.

பிரிவு இடைவெளி	அலைவெண் f	$d = \frac{x - 50{\cdot}5}{8}$	fd	fd^2
14·5—22·5	1	— 4	— 4	16
22·5—30·5	1	— 3	— 3	9
30·5—38·5	2	— 2	— 4	8
38·5—46·5	7	— 1	— 7	7
46·5—54·5	13	0	0	0
54·5—62·5	—	1	0	0
62·5—70·5	—	2	0	0
70·5—78·5	1	3	3	9
	25		—15	49

இதிலிருந்து,

$$\text{சராசரி} = \bar{x} = \frac{\Sigma f . x}{N} = \frac{1182{\cdot}5}{25} = 47{\cdot}3.$$

$$\text{மாறுபாடு} = \sigma^2 = \frac{\Sigma fx^2}{N} - (\bar{x})^2$$

$$= c^2 \left[\frac{\Sigma fd^2}{N} - \left(\frac{\Sigma fd}{N}\right)^2\right]$$

$$= (8)^2 \left[\frac{49}{25} - \left(-\frac{-15}{25}\right)^2\right]$$

$$\therefore \sigma^2 = 64\,[1{\cdot}96 - 0{\cdot}36] = 102{\cdot}4$$

எனவே, $\sigma = \sqrt{102{\cdot}4} = 10{\cdot}1$

இவற்றை μ, σ^2 உடன் ஒப்பிடலாம்.

ராண்டம் மாதிரிமுறை (ஈடுசெய்யப்படாத முறை)

ராண்டம் மாதிரி முறையினில் கண்டறியப்படும் சில உண்மைகளைக் கணக்கியல் அடிப்படையில்(mathematical basis) நிறுவுவதன்மூலமாக, அதன் இயல்பினைப்பற்றி நன்கு தெரிந்து கொள்ள முடியும்.

பொதுவாக, முழுமைத் தொகுதியின் ஒவ்வொரு கூறும் தேர்ந்தெடுக்கப்பட ஒரே வாய்ப்பு பெற்றதாக இருப்பின், அது எளிய சமவாய்ப்பு மாதிரி முறை ஆகும். வரம்புக்கு உட்பட்ட (finite) N கூறுகள் (அல்லது நபர்கள்) கொண்ட முழுமைத் தொகுதியிலிருந்து, n அலகுகள் கொண்ட ஒரு ராண்டம் மாதிரியினைத் தெரிவு செய்யவேண்டும் எனில், இவ்வாறு மாதிரிகளைத் தெரிந்தெடுக்கும் வெவ்வேறு வழிகளும் அதிகமாக அமையும். அதாவது, n—அளவுள்ள வெவ்வேறு மாதிரிகளை, N உறுப்பு முழுமைத் தொகுதியிலிருந்து,

$$\binom{N}{n} = \frac{N!}{n!\,(N-n)!}$$

வழிகளில் பெறலாம். $\binom{N}{n}$ என்பது N பொருள்களிலிருந்து n உறுப்பினை எடுக்கும் சேர்வுகளைக் (combination) குறிக்கும். [உதாரணமாக, $\binom{5}{2}$ என்பது, ஐந்து பொருட்களிலிருந்து இரண்டிரண்டாக எடுக்கப்படும் எல்லா வழிச் சேர்வுகளையும் குறிக்கும்.] மேலும் $N!$ என்பது N என்ற எண்ணின் காரணித் தொகுப்பை (factorial) குறிக்கிறது. அதாவது, 1 முதல் N வரை உள்ள முழு எண்களின் பெருக்கற் பலனைக் குறிக்கும்.

அதாவது $N! = 1.2 \ldots\ldots N$ ஆகும்.

மேற் குறிப்பிட்ட உதாரணத்தில், 5 உறுப்புகளிலிருந்து இரண்டு உறுப்புகள் கொண்ட மாதிரியானது

$$\binom{5}{2} = \frac{5!}{2!\;3!} = 10$$

வழிகளில் எடுக்கப்படலாம். A, B. C, D, E, எனும் ஐந்து எண்களிலிருந்து இரு எண் கொண்ட மாதிரிகள்

AB	AD	BC	BE	CE
AC	AE	BD	CD	DE

என்பவையாகும். ஒரே எழுத்து ஒரு மாதிரியில் ஒரு முறைதான் வரவேண்டும்.

ஒரு மாதிரியில் தெரிவு செய்யப்படும் கூறுகளைச் சமவாய்ப்பு அடிப்படையில் எவ்வாறு எடுக்கிறோம் என்பதிலும், ஒரே அளவினதான பல்வேறு மாதிரி அடைவுகளும் எவ்வாறு சமவாய்ப்புப் பெறுகின்றன என்பதிலும்தான் 'ராண்டம்' மாதிரி முறையின் உட்பொருள் அடங்குகிறது. மேற்குறிப்பிட்டபடி, தெரிந்தெடுக்கப்படும் $\binom{N}{n}$ மாதிரிகளுள் ஒவ்வொன்றும் தெரிவு செய்யப்பட சமவாய்ப்புப் பெற்றிருத்தல், சாதாரண ராண்டம் மாதிரி முறையின் (simple random sampling) அடிப்படைத் தத்துவமாகும்.

முழுமைத் தொகுதிக் கூறுகளை 1,2.........N என இலக்கமிட்டு, ராண்டம் எண் பட்டியல் அடிப்படையில் அல்லது வேறு எந்த வழியிலாவது, n அளவு கொண்ட மாதிரி தெரிவுசெய்யப்படும். $\binom{N}{n}$ மாதிரிகளும் ஒரே அளவில் சமவாய்ப்புப் பெற்றிருப்பதை எளிதில் அறியலாம்.

இதில் ஒருமுறை எடுக்கப்பட்ட கூறுனது (unit) ஈடு செய்யப்படவேண்டிய (replacement) அவசியமில்லை. ஏனெனில், முன்பு எடுக்கப்பட்ட (கூறு) அலகு மீண்டும் வரநேரிடும். எனவே, இத்தகைய வழிமுறையை, ஈடுசெய்யப்படாத, மாதிரி முறை (sampling without replacement) என அழைக்கலாம். N மிகப் பெரிய எண்ணுகை இருப்பதால், ஒவ்வொரு கூறும் சம அளவில், தெரிந்தெடுக்கப்படும் வாய்ப்புப் பெற்றிருக்கும். ஈடு செய்யப்படும் முறையில் (with replacement), மாதிரித் தெரிவு முறை எளிதாக அமைவதால், சில நிலைகளில் அதுவும் பயன் படுத்தப்படுகிறது.

$N = 300$ என்றபடி உள்ள முழுமைத் தொகுதியிலிருந்து, மாதிரிக் கூறினை எடுக்கும்போது, முதலில் தெரிவு செய்யும் கூறுனது (unit), $\frac{1}{300}$ எனும் தேர்வு வாய்ப்பை (chance or probability) பெற்றிருக்கும். தேர்வு செய்யப்படும் கூறினை (உறுப்பை) நீக்கிவிட்டு, அடுத்த கூறினைத் தெரிந்தெடுக்கையில் அதற்கு $\frac{1}{299}$ என்றபடியான வாய்ப்பு இருக்கக் காணலாம்.

இதே போல், தேவையான அளவினதான (size) மாதிரிக் கூறு தெரிந்தெடுக்கப்படும் முறையானது ஈடு செய்யப்படாத முறை (without replacement) எனப்படும். இதில் கிடைக்கும் மாதிரி

யானது 'எளிய' ராண்டம் மாதிரிக்கூறு (simple random sample) அல்ல. மாருக, ஒருமுறை எடுக்கப்பட்ட கூறு, மீண்டும் சேர்த்துக்கொள்ளப்பட்டு, ஈடு செய்யப்படும் (with replacement) முறையில், ஒவ்வொரு கூறின் தெரிவும், $\frac{1}{300}$ என்ற சம வாய்ப்பு பெறுவதை அறியலாம்.

மாதிரி மதிப்பீடுகள் (Sample Estimates)

மாதிரி முறை கணிப்பாய்வுகளில் (Sampling surveys), மாதிரியில் சேர்க்கப்படும் ஒவ்வொரு கூறினுக்கும் அல்லது அலகுக்கும் சில இயல்புகளை (அளவைகளை) குறித்து வைக்க முற்படுவோம். இவற்றைப் பொதுவாக 'சிறப்பு இயல்புகள்' (characteristics) என்று குறிப்பிடுவோம்.

N அளவு கொண்ட (size) முழுமைத் தொகுதியின் அளந்தறியப்படும் மதிப்புக்களை y_1, y_2 ....y_N என்று குறிப்பிடலாம். y_i என்பது மாறும் சிறப்பினை (varying characteristic) குறிக்கும் மாறி (variable) ஆகும். இவற்றின் தொடர்பான மாதிரி மதிப்புகள் y_1, y_2 .....y_n என்று குறிக்கப்பெறும். i—எனும் மாதிரிக்கூறு மதிப்பு y_i என்று ஆக, $i = 1, 2$ ...... n என்று மதிப்புகள் கொடுக்கப்படலாம். y_1, y_2 ..... என்னும் மதிப்புகள், வாய்ப்பு அடிப்படையில் பெறப்படுபவை என்பது குறிப்பிடத் தக்கது.

பெரிய எழுத்துக்கள் முழுமைத் தொகுதிக்கும், சிறிய எழுத்துக்கள் மாதிரிக் கூற்றுக்கும் பயன்படுத்தப்படும்.

கூட்டு மொத்தம், கூட்டுச் சராசரி ஆகியவை பின்வருமாறு குறிக்கப்படும் :

முழுமைத் தொகுதி : மொத்தம் : $Y = Y_1 + Y_2 + \ldots\ldots + Y_n$

சராசரி : $\bar{Y} = \frac{Y_1 + Y_2 + \ldots\ldots + Y_n}{N}$

$$= \frac{Y}{N}$$

மாதிரி : மொத்தம் : $y = y_1 + \ldots\ldots\ldots + y_n$

சராசரி : $\bar{y}$: $\frac{y_1 + \ldots\ldots\ldots + y_n}{n}$

$$: \frac{y}{n}$$

மாதிரி முறையானது மூன்றுவிதமாக முக்கியத்துவம் உடைய தாகக் கருதப்படும். முதலாவதாக, ஏதேனும் ஒரு சிறப்பினை பற்றி, எல்லாக் கூறுகளையும் கருதி, பெறப்படும் மதிப்பு அளவு களின் கூட்டு மொத்தம் (Y) அறியப்படுவது.

உதாரணமாக தேர்ந்தெடுக்கப்படும் ஓர் உணவுப் பயிரானது (food-crop) விளைவிக்கப்படும் நிலப்பரப்பின் மொத்தத்தை (ஏக்கர் அளவில்) ஏதெனும் ஓர் இடத்திற்கு மதிப்பிடுவதாகக் கொள்ளலாம். இரண்டாவதாக, கருதப்படும் கூறுபற்றிய (unit) சராசரி அளவினை மதிப்பிடுவதாகும். உதாரணமாக, பயிர் விளைச்சல் நிலத்தின் சராசரிப் பரப்பின் அளவைக் குறிப்பிடலாம். மூன்ருவதாக, ஒரு குறிப்பிட்ட பிரிவில் (class) இடம்பெறும் அலகுகளின் விகித சம (proportion) அளவு பெறப்படுவது. அதாவது, மொத்த நிலப்பரப்பில், குறிப்பிட்ட பயிரானது பயிர் செய்யப் படாத நிலப்பரப்பினது விகித மதிப்பு எனக் கொள்ளலாம்

Y என்ற மதிப்பின் (value) மாதிரி மதிப்பீடானது (estimate) $\hat{Y}$ என்று குறியிடப்பெறும்.

மதிப்பீடுகளின் சில தன்மைகள் (Properties)

n அளவுடைய பல்வேறு மாதிரிகளையும் ஒரு முழுமைத் தொகுதியினின்று கருதி. ஏதேனும் ஒரு மதிப்பீட்டு அளவின் (estimate) சராசரியினை (mean) கணக்கிட்டு, அச் சராசரி யானது முழுமைத் தொகுதியின் உண்மை அளவை மதிப்பாக (Population parameter) அமைந்திருப்பின், மதிப்பிடும் முறை யினைப் பிறழ்ச்சியற்றது (unbiased) என்று கூறுகிறோம். அதாவது, θ என்ற முழுமைத் தொகுதி அளவையானது (statistic)

$$E(t) = \theta$$

என்ற விதிக்கு உட்படின், t என்ற மதிப்பீடு பிறழ்ச்சியற்றதாக (unbiased) ஆகும், E என்பது எல்லா மாதிரிகளையும் கருதிப் பெறப்படும் சராசரியினைக் குறிக்கும்.

தேற்றம்: 1 மாதிரிச் சராசரி (sample mean) $\bar{y}$ ஆனது முழுமைத் தொகுதிச் சராசரியின் பிறழ்ச்சியற்ற (unbiased) மதிப்பீடாகும்.

நிருபணம்: மேற்கண்ட வரையறைப்படி, $E(\bar{y}) = \bar{Y}$ என நிறுவவேண்டும்.

$$\therefore \quad E(\bar{y}) = \frac{\Sigma \bar{y}}{\binom{N}{n}} = \frac{\Sigma (y_1 + \ldots + y_n)}{n \dfrac{N!}{n!\,(N-n)!}} \qquad (1.1)$$

கூட்டுத் தொகை, எல்லா $\binom{N}{n}$ மாதிரிகளையும் கருதிய தாகும். இக் கூட்டுத் தொகையை மதிப்பிட, ஏதேனும் ஒரு y_i மதிப்பு எத்தனை மாதிரிகளில் வரும் என்று காணவேண்டும். எஞ்சியுள்ள மாதிரிகட்கு $(N-1)$ கூறுகளும், மாதிரியில் மீதம் $(n-1)$ இடங்களும் அமைவதால், y_i-மதிப்பு இடம்பெறும் மொத்த மாதிரிகள்

$$\binom{N-1}{n-1} = \frac{(N-1)!}{(n-1)!\;(N-n)!}$$

என்று கணக்கிடப்படும்.

எனவே,

$$\Sigma (y_1 + y_2 + \ldots\ldots + y_n) = \frac{(N-1)!}{(n-1)!\,(N-n)!}(Y_1 + \ldots \ldots + Y_N)$$

ஆகிறது. $\left\{ \because \binom{N-1}{n-1} Y_1 + \binom{N-1}{n-1} Y_2 + \ldots\ldots + \binom{N-1}{n-1} Y_N = \binom{N-1}{n-1} (Y_1 + \ldots\ldots + Y_N) \right\}$

இதை (1·1)-ல் பிரதியிட

$$E(\bar{y}) = \frac{(N-1)!}{(n-1)!\,(N-n)!} \cdot \frac{n!\,(N-n)!}{nN!}(Y_1 + \ldots \ldots + Y_N)$$

$$= \frac{(Y_1 + \ldots\ldots + Y_N)}{N}$$

$$\therefore \quad E(\bar{y}) = \bar{Y}$$

என்று நிறுவப்படும்.

(ஏனெனில் $\dfrac{n!}{(n-1)!} = n$ என்பது சூத்திரம்.)

துணை முடிவு (Corallary)

மேற்கண்ட உண்மையிலிருந்து, $\hat{Y} = N \cdot \bar{y}$ என்பது முழுமைத் தொகுதி மொத்த மதிப்பின் (total), அதாவது Y-ன் பிறழ்ச்சியற்ற மதிப்பீடாகப் பெறப்படுகிறது.

மேற்குறித்த $\hat{Y}$ என்ற மதிப்பீட்டில்,

$$\text{அதாவது} \quad \hat{Y} = N \cdot \bar{y} = N \cdot \left(\frac{\sum_{i=1}^{n} y_i}{n} \right)$$

என்ற வரையறையில், $\frac{N}{n}$ என்ற காரணி அளவானது (factor), வீக்கக் காரணி அல்லது பெருக்கற் காரணி ((inflation factor) என்று கூறப்படும். அதன் தலைகீழ் மதிப்பு (reciprocal), அதாவது $\frac{n}{N}$ என்றும் அளவு, மாதிரியின் அளவிற்கும் (sample size), முழுமைத் தொகுதி அளவிற்கும் (population size) இடையில் உள்ள விகித அளவாக, மாதிரிப் பின்னம் (Sampling fraction) என்று கூறப்படும்.

$E(\bar{y}) = \bar{Y}$ என்ற முடிவிலிருந்து,

$$E \frac{(y_1 + \ldots\ldots + y_n)}{n} = \frac{Y_1 + \ldots\ldots + Y_N}{N}, \text{ அதாவது}$$

$$E(y_1 + \ldots\ldots + y_n) = \frac{n}{N}(Y_1 + Y_2 + \ldots\ldots + Y_N) \text{ எனப்}$$

பெறலாம்.

வரம்புடைய முழுமைத் தொகுதித் திருத்தம் (Finite Population Correction)

எண்ணிக்கை வரம்பற்ற (infinite) ஒரு முழுமைத் தொகுதியினின்று, n-அளவுள்ள ஒரு சம வாய்ப்பு மாதிரி (random sample) எடுக்கப்படுகையில், அம் மாதிரியின் சராசரி அளவினது மாறுபாட்டு மதிப்பு (Variance of mean) $\frac{\sigma^2}{n}$ என நிறுவப்படுகிறது. அதாவது, $\bar{y}$ என்ற மாதிரிச் சராசரியின், மாதிரி முறை சார் மாறுபாட்டு அளவு (sampling variance).

$$V(\bar{y}) = E\{\bar{y} - E(\bar{y})\}^2$$

$$= E\left\{\frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{Y})\right\}^2 \quad (\because E(\bar{y}) = \bar{Y})$$

$$= \frac{1}{n^2} \sum_{i=1}^{n} E(y_i - \bar{Y})^2$$

$$+ \frac{1}{n^2} \sum_{i=1}^{n} \sum_{i \neq j} E(y_i - \bar{Y})(y_j - \bar{Y})$$

$\therefore \quad V(\bar{y}) = \frac{\sigma^2}{n}$ ஆகும்.

ஏனெனில், $E(y_i - \bar{Y})^2 = V(y_i) = \sigma^2$

மேலும் $E(y_i - \bar{Y})(y_j - \bar{Y})$ என்பது y_i, y_j இவற்றின் உடன் மாறுபாட்டு அளவை (covariance) குறிக்கையில், இதன் மதிப்பு பூச்சியமாகும் $(i \neq j)$. $\because$ i-ஆவது கூறினை தெரிவு செய்வதும், j-ஆவது கூறினை தெரிவு செய்வதும் ஒன்றோடொன்று முற்றிலும் தொடர்புடையன அல்ல (independent).

ஆனால் முற்றும் மாறாக, முழுமைத் தொகுதியானது எண்ணிக்கை வரம்புடையதாக (finite) இருப்பின் $\left(\frac{N-n}{N}\right)$ என்ற காரணி (factor) புதிதாகச் சேர்க்கப்படுவது, மேற்கண்ட முடிவினின்று ஏற்படும் மாற்றமாக அமைகிறது. இக் காரணியானது, வரம்புடை முழுமைத் தொகுதி திருத்தக் காரணி எனப்பட்டு, சுருக்கமாக '*fpc*' என்று குறிக்கப் பெறும். மாதிரி எடுத்தல் பின்ன அளவான (sampling fraction), $\frac{n}{N}$ என்ற அளவு மிகக் குறைவாக (low) இருப்பின், '*fpc*' காரணி, '1' எனும் மதிப்புக்கு ஏறக்குறைய சமமாகிறது. அப்பொழுது, முழுமைத் தொகுதி அளவானது (size), (அதாவது, N), மாதிரிச் சராசரியின் தரப்பிழையின் அளவை (standard error) எந்த விதத்திலும் நேரடியாகப் பாதிப்பதில்லை. நடைமுறையில், மாதிரி எடுத்தல் பின்னமானது 5%-க்கு கீழிருந்தால், *fpc* பொருட்படுத்தப்படுவதில்லை.

மதிப்பீடுகளின் மாறுபாட்டு அளவுகள்: (Variances of Estimates)

வரம்புள்ள ஒரு முழுமைத் தொகுதியின் y_i என்ற மதிப்பின் மாறுபாடானது (variance)

$$\sigma^2 = \frac{\sum_{i=1}^{N} (Y_i - \bar{Y})^2}{N}$$

என்று வரையறுக்கப்படும். இதே முழுமைத் தொகுதிக்கு,

$$S^2 = \frac{\sum_{i=1}^{N} (Y_i - \bar{Y})^2}{N-1}$$

என்றும் ஒரு மாறுபாட்டு அளவினை வரையறுப்போம். இதன் மூலமாக, மற்றொரு உண்மையினை (property) உறுதி செய்ய முடிகிறது.

தேற்றம் 2

ஓர் எளிய ராண்டம் மாதிரிக் கூறின் (simple random sample) சராசரியான $\bar{y}$ என்பதன் மாறுபாட்டு அளவு (variance).

$$V(\bar{y}) = E(\bar{y} - \bar{Y})^2$$
$$= \frac{S^2}{n} \cdot \frac{N-n}{N}$$
$$= \frac{S^2}{n} \cdot (1-f)$$

என்பதாகும். மாதிரிச் சராசரியின் மாறுபாடு (Variance of sample mean) என்று இதனைக் கூறுகிறோம்.

நிறுவுதல் (Proof)

$$n \cdot \bar{y} = (y_1 + \ldots\ldots + y_n)$$

என்பதால்,

$$n(\bar{y} - \bar{Y}) = (y_1 - \bar{Y}) + (y_2 - \bar{Y}) + \ldots\ldots + (y_n - \bar{Y})$$

என்று எழுதலாம். ...... (2·1)

எனவே,

$$E(y_1 + \ldots\ldots + y_n) = \frac{n}{N}(Y_1 + \ldots\ldots + Y_N)$$

என்ற உண்மையைப் பயன்படுத்தி,

$$E[(y_1 - \bar{Y})^2 + (y_2 - \bar{Y})^2 + \ldots\ldots + (y_n - \bar{Y})^2]$$

$$\frac{n}{N}[(y_1 - \bar{Y})^2 + \ldots\ldots + (Y_N - \bar{Y})^2]$$

என எழுதலாம். ...... (2·2)

மேலும்,

$$E[(y_1 - \bar{Y})(y_2 - \bar{Y}) + (y_2 - \bar{Y})(y_3 - \bar{Y}) + \ldots\ldots + (y_{n-1} - \bar{Y})(y_n - \bar{Y})]$$

$$= \frac{n(n-1)}{N(N-1)}[(Y_1 - \bar{Y})(Y_2 - \bar{Y}) + \ldots \ldots + (Y_{N-1} - \bar{Y})(Y_N - \bar{Y}) \quad \ldots\ldots (2\cdot3)$$

என்றும் ஆகும்.

(2·1)-ல் உள்ளதன் வர்க்கத்தைக் கண்டு, எல்லா மாதிரிகளின் அடிப்படையில் சராசரி எடுக்க, வர்க்கச் சராசரியினை

$$n^2 \cdot E(\bar{y} - \bar{Y})^2 = E\{(y_1 - \bar{Y})^2 + (y_2 - \bar{Y})^2 + \ldots\ldots + (y_n - \bar{Y})^2 + 2[(y_1 - \bar{Y})(y_2 - \bar{Y}) + \ldots\ldots + (y_{n-1} - \bar{Y})(y_n - \bar{Y})]\}$$

$$= \frac{n}{N}\left\{(y_1 - \bar{Y})^2 + \ldots\ldots + (y_N - \bar{Y})^2 + 2\cdot\frac{(n-1)}{(N-1)}[(Y_1 - \bar{Y})(Y_2 - \bar{Y}) + \ldots\ldots + (Y_{N-1} - \bar{Y})(Y_N - \bar{Y})]\right\} \quad [(2\cdot3)\text{-ன் மூலம்}]$$

$$= \frac{n}{N}\left\{\left(1 - \frac{n-1}{N-1}\right)[(Y_1 - \bar{Y})^2 + \ldots\ldots + (Y_N - \bar{Y})^2] + \left(\frac{n-1}{N-1}\right)[(Y_1 - \bar{Y}) + \ldots\ldots + (Y_N - \bar{Y})]^2\right\} \quad \ldots\ldots (2\cdot4)$$

(2·4)-வலது புறம் உள்ள தொடரில் (expression), இரட்டை அடைப்பினுள் இரண்டாவது உறுப்பின் (term) மதிப்பு பூச்சியமாகும். ஏனெனில், $Y_1 + Y_2 + \ldots\ldots + Y_N = N \cdot \bar{Y}$ ஆகும். பிறகு, (2·4)-ன் இருபுறமும், n^2-ஆல் வகுக்க,

$$E(\bar{y} - \bar{Y})^2 = \frac{(N-n)}{n \cdot N(N-1)} \cdot [(Y_1 - \bar{Y})^2 + \ldots\ldots + (Y_n - \bar{Y})^2]$$

$$= \frac{(N-n)}{n \cdot N(N-1)} \sum_{i=1}^{N} (Y_i - \bar{Y})^2.$$

ஆனால், வரையறைப்படி,

$$V(\bar{y}) = E(\bar{y} - \bar{Y})^2 = \frac{(N-n)}{N} \cdot \frac{S^2}{n} \text{ என வரும்.}$$

$$\left\{ \because \; S^2 = \frac{1}{N-1} \sum_{i=1}^{N} (Y_i - \bar{Y})^2 \right\}.$$

$$\text{எனவே } S(\bar{y}) = \frac{S^2}{n} \cdot \frac{(N-n)}{N} \text{ என நிறுவப்படுகிறது.}$$

துணைமுடிவு (Corallary) (1):

மேற்கண்ட தேற்றத்திலிருந்து, $\bar{y}$ என்ற மாதிரி அளவையின் (statistic) தர விலக்க அல்லது தரப் பிழையின் (standard error) மதிப்பை

$$\sigma_{\bar{y}} = \frac{s}{\sqrt{n}} \cdot \sqrt{\frac{N-n}{N}} = \frac{s}{\sqrt{n}} \sqrt{1-f}$$

என்று பெறுகிறோம்.

{ ஒரு மாதிரி அளவையின் மாறுபாட்டு மதிப்பினது வர்க்க மூலத்தை, அம் மாதிரி அளவையினது தரவிலக்க அளவாகக் கருதப்படும்.}

துணை முடிவு (2)

முழுமைத் தொகுதி கூட்டு மொத்தமான 'Y' எனும் மதிப்பின் மாதிரி மதிப்பீடான $\hat{Y} = N.\ \bar{y}$ என்பதனுடைய மாறுபாட்டு அளவு (variance)

$$(\hat{Y}) = E\left(\hat{Y} - \bar{\hat{Y}}\right)^2$$
$$= E(N\bar{y} - N.Y)^2$$
$$\{\because E(N.\bar{y}) = N.\bar{Y}\}$$
$$= N^2.\ E(Y - \bar{Y})^2$$
$$= N^2.\ V(\bar{y})$$

$$= N^2 . \frac{s^2}{n} . \frac{(N-n)}{N}$$

அதாவது, $V(\hat{Y}) = \frac{N^2 . s^2}{n} . (1-f)$

என்று பெறப்படும்.

துணை முடிவு (3)

துணை முடிவு (2)ன் அடிப்படையில், $\hat{Y}$ எனும் மதிப்பீட்டின் த்ரவிலக்க அளவு (standard error)

$$\sigma_{\hat{y}} = \frac{NS}{\sqrt{n}} . \sqrt{(1-f)}$$

ஆகிறது.

இம் மூன்று முடிவுகளும் அதிக அளவில் பயன்படுத்தப் படுகின்றன.

தேற்றம் 3

ஒரு எளிய ராண்டம் மாதிரியின் மாறுபாடு

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2$$

என்பது,

$$S^2 = \frac{1}{N-1} \sum_{i=1}^{N} (Y_i - \bar{Y})^2$$

என்ற அதன் முழுமைத் தொகுதி மாறுபாட்டின் பிறழ்ச்சியற்ற மதிப்பீடாகும் (Unbiased estimate),

நிருபணம் :

மாதிரி மாறுபாட்டைப் பின்வருமாறு எழுதுவோம்:

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2$$

$$= \frac{1}{n-1} \sum_{i=1}^{n} \left[(y_i - \bar{Y}) - (\bar{y} - \bar{Y}) \right]^2$$

$$= \frac{1}{n-1}\left[\sum_{i=1}^{n}(y_i-\bar{Y})^2 - n(\bar{y}-\bar{Y})^2\right]$$

[ஏனெனில், $2\sum_{i=1}^{n}(y_i-\bar{Y})(\bar{y}-\bar{Y})$

$$= 2(\bar{y}-\bar{Y})\sum_{i=1}^{n}(y_i-\bar{Y})$$

$$= 2(\bar{y}-\bar{Y})(n\bar{y}-n\bar{Y})$$

$$= 2n(\bar{y}-\bar{Y})^2$$

என ஆகிறது.]

எல்லா n- அளவுள்ள மாதிரிகளையும் கருதி, சராசரி எடுக்க,

$$E\left[\sum_{i=1}^{n}y_i\right] = \frac{n}{N}\left[\sum_{i=1}^{n}Y_i\right]$$

என்ற விதிப்படி,

$$E\left[\sum_{i=1}^{n}(y_i-\bar{Y})^2\right] = \frac{n}{N}\left[\sum_{i=1}^{N}(Y_i-\bar{Y})^2\right]$$

$$= \frac{n.\ (N-1)S^2}{N}$$

என வரும் $\left[\because s^2 = \frac{1}{N-1}\sum_{i=1}^{N}(Y_i-\bar{Y})^2\right]$

மேலும், சென்ற தேற்றத்தைப் பயன்படுத்தி,

$$E\ [n(\bar{y}-\bar{Y})^2] = n.\ E(\bar{y}-\bar{Y})^2$$

$$= n.\ V(\bar{y})$$

$$= n.\ \frac{(N-n)}{N}.\ \frac{S^2}{n}$$

$$= \frac{(N-n)}{N}.\ S^2$$ என்றும் பெறலாம்.

$$\therefore \quad E(s^2) = \frac{1}{(n-1)}$$

$$\left[E \sum_{i=1}^{n} (y_i - \bar{Y})^2 - E\left(n(\bar{y}-\bar{Y})^2 \right) \right]$$

$$\{ \because \quad E(x+y) = E(x) + E(y) \}$$

$$= \frac{S^2}{(n-1).N} [\, n(N-1) - (N-n) \,]$$

$$= \frac{S^2}{(n-1).N} \cdot N(n-1)$$

$$= S^2$$

அதாவது $E(s^2) = S^2$ என்று நிறுவப்படுகிறது,

துணை முடிவுகள்

மேற்கண்ட தேற்றத்திலிருந்து,

(i) $\bar{y}$ என்ற மாதிரிச் சராசரியின் மாறுபாடு,

$$v(\bar{y}) = s^2_{\bar{y}} = \frac{s^2}{n}\left(\frac{N-n}{n}\right) = \frac{s^2}{n}(1-f)$$

என்ற பிறழ்ச்சியற்ற மதிப்பீட்டைப் (Unbiased estimate) பெறும்.

(ii) இதே போல, $\hat{Y} = N\bar{y}$ என்ற மொத்த மதிப்பீடானது

$$v(\hat{Y}) = s^2_{\hat{Y}} = \frac{N^2.s^2}{n}\left(\frac{N-n}{n}\right) = \frac{N^2.s^2}{n}(1-f)$$

என்ற பிறழ்ச்சியற்ற மதிப்பீட்டைப் பெறும்.

{ குறிப்பு : உண்மை மதிப்பு $\sigma^2_{\bar{y}} = V(\bar{y})$

எனவும், மாதிரிமூலமாக மதிப்பிடப்பட்ட மதிப்பு

$s^2_{\bar{y}} = v(\bar{y})$

எனவும் குறிக்கப்படுவது, கவனிக்கத்தக்கது. }

இனி, 'எளிய ராண்டம் முறை' எனப்படும், 'ஈடு செய்யப் படும் முறை' (With Replacement) பற்றிய சில முடிவுகளை ஆராய்வோம்:

இரு உறுப்பு மாதிரி முறை (ஈடு செய்யப்படும் முறை)
[Sampling of two Units) (With replacement method]

N உறுப்புகள் கொண்ட முழுமைத் தொகுதியிலிருந்து இரண்டு உறுப்புகளுள்ள (units) உள்ள எளிய ராண்டம்

மாதிரிகள் எடுக்கப்படுவதைக் கருதுவோம். இங்கு முதலில். தெரிவு செய்யப்பட்ட உறுப்பினை அல்லது கூறினை, ஈடுசெய்து விட்டு, இரண்டாவது உறுப்பினை அதே முறையில் அதே முழுமைத் தொகுதியிலிருந்து தெரிவு செய்கிறோம். அதாவது, y_1, y_2 என்பவை முதலிலும், பிறகே தொடர்ந்து இரண்டாவதாகவும் எடுக்கப்பட்ட மதிப்புகளாக (Values) கொள்வோம். y_1-ஐத் தெரிவு செய்யப்பட்ட வழிமுறையில் y_2 என்ற மதிப்பும் தெரிவு செய்யப்படுவதால், y_2 மதிப்பும் $\bar{Y}$-ன் பிறழ்ச்சியற்ற (Unbiased) மதிப்பீடாகும். அதனுடைய (y_2) மாறுபாடும் σ^2 ஆகும். எனவே, இவ்விரண்டு மதிப்புக்களின் சராசரி, $\bar{y} = \frac{y_1 + y_2}{2}$, $\bar{Y}$-ன் பிறழ்ச்சியற்ற மதிப்பீடாக அமையும். இரண்டு கூறுகளும் ஒன்றோடொன்று சார்பற்றபடி (Independent) தெரிவு செய்யப்படுவதால், y_1, y_2 இவற்றின் உடன் மாறுபாட்டு (Covariance) அளவு பூச்சியமாக [அ—து, $\text{cov}(y_1, y_2) = 0$] $\bar{y}$—ன் மாதிரி முறைசார் மாறுபாட்டு அளவு (Sampling Variance) $V(\bar{y}) = \frac{1}{4}[V(y_1) + V(y_2)] = \frac{\sigma^2}{2}$ என்று கணிக்கப்படும்.

எனவே, மாதிரியில் உள்ள கூறுகள் (உறுப்புகள்) எண்ணிக்கை, அதிகரிக்கையில், மாதிரி முறைசார் மாறுபாட்டு அளவு குறையக் காணலாம். அதாவது, மாதிரி அளவு (Sample Size) அதிகரிக்க அதிகரிக்க, குறிப்பிட்ட திட்பத்துடன் கூடிய (Precision) மதிப்பீடுகளைப் பெற ஏதுவாகிறது.

பொதுவாக ஈடு செய்யப்படும் சாதாரண ராண்டம் மாதிரி முறையானது (SRSWR) N அளவுள்ள முழுமைத் தொகுதியிலிருந்து, ஒவ்வொரு உறுப்பின் தெரிவிலும் முந்தைய உறுப்பையும் ஈடுசெய்துவிட்டு சமவாய்ப்புடன் n—அலகுகளைத் தேர்ந்தெடுப்பதாகும் என வரையறுக்கலாம். ஒருமுறை எடுத்த கூறு திரும்பவும் வர வாய்ப்பு உண்டு.

ஈடுசெய்யப்படும் முறையில், n—அளவு எளிய ராண்டம் மாதிரித் தெரிவில் சில முடிவுகள் (SRSWR)

மாதிரி சராசரியின் எதிர்பார்த்த மதிப்பு (Expected Value of Sample Mean)

$Y_1, Y_2 \ldots\ldots Y_N$ என்ற முழுமைத் தொகுதிக் கூறுகளிலிருந்து, n—கூறுகள் உள்ள ($y_1, y_2 \ldots\ldots y_n$) எனும் மாதிரி, குறிப்பிட்ட

முறைப்படி எடுக்கப்படுகிறது. i—வது நிலையில் எடுக்கப்படும் y_i எனும் மதிப்பானது, வாய்ப்பு மாறி (random variable) ஆகக் கருதப்பட்டு, அதன் மதிப்புகள் $Y_1, Y_2 ...... Y_N$ என்பவை

சமவாய்ப்பில் அமையும். மாதிரிச் சராசரியான $\bar{y} = \frac{1}{n} \sum_{i=1}^{n} y_i$

என்பது $\bar{Y}$—ன் பிறழ்ச்சியற்ற மதிப்பாகும்.

$$\text{அதாவது } E(\bar{y}) = \frac{1}{n} \sum_{i=1}^{n} E(y_i) = \bar{Y}$$

ஒரே மதிப்பு பலமுறை மீண்டும் தெரிந்தெடுக்கப்படலாம் என்பது கவனிக்கத்தக்கது.

மாதிரி சராசரியின் மாறுபாடு :

$\bar{y}$—ன் மாதிரிமுறை சார்ந்த மாறுபாட்டு அளவு (Sampling Variance)

$$V(\bar{y}) = E[\bar{y} - E(\bar{y})]^2$$

$$= E\left[\frac{1}{n} \sum_{i=1}^{n} (y_i - \bar{Y})^2\right]$$

$$= \frac{1}{n^2} \sum_{1}^{n} E(y_i - \bar{Y})^2$$

$$+ \frac{1}{n^2} \sum_{i=1}^{n} \sum_{j \neq i} E(y_i - \bar{Y})(y_i - \bar{Y})$$

$$= \frac{\sigma^2}{n}$$

ஏனெனில் $E(y_i - \bar{Y})^2 = \sigma^2$ எனவும்

$E(y_i - \bar{Y})(y_j - \bar{Y}) = 0$ எனவும் பெறப்படும் (i, j ஆகியவை தொடர்பற்றவை ஆதலால், y_i, y_j ஆகியவற்றின் உடன்வேறுபாடு பூச்சியமாகும்.)

$\bar{y}$—ன் தரப்பிழை அளவு $\sigma_{\bar{y}} = \frac{\sigma}{\sqrt{n}}$ ஆகும்.

$\bar{y}$-ன் மாறுபாட்டுக் கெழு (coefficient of variation),

$$C\,(\bar{y}) = \frac{1}{\sqrt{n}}\,\frac{\sigma}{\bar{Y}} \text{ ஆகும்.}$$

n—அதிகரிக்கையில், ஒரு கட்டத்திற்குப் பின், தரப்பிழையின் அளவானது பொருட்படுத்தப்படாத அளவு குறையக் காணலாம்.

$\bar{y}$—ன் மாதிரி முறை சார்ந்த மாறுபாட்டுக்கு பிறழ்ச்சியற்ற மதிப்பீடு :

மாதிரி அளவையின் எதிர்பார்த்த மதிப்பினைக் கருதுவதன் மூலம் (Expected Value), பிறழ்ச்சியற்ற மதிப்பீடு பெறப் படுகிறது. இங்கு மாதிரி அளவையான (Statistic)

$$\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2$$

என்பதைக் கருதுவோம். இதன் எதிர்பார்த்த மதிப்பு

$$E\left[\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2\right] = E\left\{\frac{1}{n}\sum_{i=1}^{n}y_i^2-(\bar{y})^2\right\}$$

$$=\frac{1}{n}\sum_{i=1}^{n}E(y_i^2) - E(\bar{y})^2$$

$$=(\sigma^2+\bar{Y}^2)-\left(\frac{\sigma^2}{n}+\bar{Y}^2\right)$$

$$=\frac{(n-1)}{n}\sigma^2$$

{ஏனெனில், $E(y_i^2) = V(y_i)+\bar{Y}^2 = \sigma^2+\bar{Y}^2$

$\therefore$ மேலும் $E(\bar{y}^2) = V(\bar{y})+\bar{Y}^2$ ஆகும் }

இதிலிருந்து, $V(\bar{y})$ அதாவது $\frac{\sigma^2}{n}$ —ன் பிறழ்ச்சியற்ற மதிப்பீடானது,

$$V(\bar{y}) = \frac{s^2}{n}\left[s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(y_i-\bar{y})^2\right]$$

என்பது பெறப்படும்.

எனவே, $\bar{y}$ - ன் தரப்பிழை அளவு

$$\hat{\sigma}_{\bar{y}} = s_{\bar{y}} = \frac{s}{\sqrt{n}}$$ எனவும்

$\bar{y}$—ன் மாறுபாட்டுக் கெழு

$$\hat{C}_{\bar{y}} = c_{\bar{y}} = \frac{1}{\sqrt{n}} \frac{s}{\bar{y}}$$ எனவும் மதிப்பீடு

(estimate) செய்யப்படும்.

ஈடு செய்யப்படும் முறையில், $E(s^2) = \sigma^2$ என்று பெறப் படும் :

வரையறை :

$$\sigma^2 = \frac{1}{N} \sum_{N} (y_i - \bar{Y})^2$$

சமவாய்ப்பு மாதிரியிலிருந்து,

$$s^2 = \frac{1}{n-1} \sum_{n} (y_i - \bar{y})^2$$

என்று அறிகிறோம்.

அதாவது $s^2 = \frac{1}{n-1} \left[\Sigma (y_i - \bar{Y})^2 - n (\bar{y} - \bar{Y})^2\right]$

இதிலிருந்து

$$E(s^2) = \frac{1}{n-1} \left\{E\left[\Sigma (y_i - \bar{Y})^2\right] - E\left[n (\bar{y} - \bar{Y})^2\right]\right.$$

முதலில்

$$E\left[\sum_{n} (y_i - \bar{Y})^2\right] = \sum_{n} E (y_i - \bar{Y})^2$$

$$= \sum_{1}^{n} \sum_{1}^{N} \frac{1}{N} (Y_i - \bar{Y})^2$$

$$= \frac{1}{N} \cdot \sum_{n} (N-1)\, s^2$$

$$\left\{ \because \quad S^2 = \frac{1}{N-1} \sum_{N} (y_i - \bar{Y})^2 \right\}$$

$$= \frac{n(N-1)}{N}\, s^2$$

மேலும்

$$E\,[n\,(\bar{y} - \bar{Y})^2] = n \cdot E\,(\bar{y} - \bar{Y})^2$$

$$= n . \frac{\sigma^2}{n} = \sigma^2 \qquad \left[\because \text{ ஈடு செய்யப்படும் முறை} \right]$$

எனவே

$$E\,(s^2) = \frac{1}{n-1} \left\{ \frac{n\,(N-1)}{N} . \ S^2 - \sigma^2 \right\}$$

$$E\,(s^2) = \frac{(n-1)\,\sigma^2}{(n-1)} = \sigma^2$$

என்று பெறப்படும்.

எடுத்துக்காட்டு 1 :

(i) முழுமைத் தொகுதி (1, 2, 3, 4) என்பதிலிருந்து, $n = 2$ என்ற அளவில் மாதிரிகளை, ஈடு செய்யப்படாத ராண்டம் மாதிரி முறைப்படி எடுத்து, பல்வேறு மதிப்பீடுகள் காணல் :

மொத்தம் $\binom{4}{2} = \frac{4 \times 3}{1 \times 2} = 6$ மாதிரிகள் எடுக்கப்படும்.

மாதிரி எண்	மதிப்புகள் y_1	y_2	சராசரி $\bar{y}$	$\bar{y} - \bar{Y}$	$(\bar{y} - \bar{Y})^2$
1	1	2	1·5	— 1·0	1·0
2	1	3	2·0	— 0·5	0·25
3	1	4	2·5	0	0
4	2	3	2·5	0	0
5	2	4	3·0	0·5	0·25
6	3	4	3·5	1·0	1·00

முழுமைத் தொகுதி சராசரி : $\bar{Y} = 2{\cdot}5$. $(N = 4)$

மாதிரி $\bar{y}$ மதிப்புக்களின் எதிர்பார்த்த அளவு $\Big\} = \frac{1}{6}\left[1 \cdot 5 + \ldots + 3 \cdot 5\right]$

$$= \frac{15}{6} = 2 \cdot 5$$

எனவே $E(\bar{y}) = Y$ (அ—து) $\bar{y}$ ஆனது $\bar{Y}$ மதிப்பின் பிறழ்ச்சியற்ற மதிப்பீடாகும்.

$\bar{y}$-ன் வேறுபாடு, முழுமைத்தொகுதி பற்றியது.

$$V(\bar{y}) = \frac{\Sigma(\bar{y} - Y)^2}{6} = \frac{2 \cdot 5}{6} = 0 \cdot 417$$

மேலும், நியமித்தபடி,

$$V(\bar{y}) = \frac{S^2}{n} \cdot \frac{N - n}{N} \text{ ஆகும்.}$$

இதில்,

$$S^2 = \frac{1}{N-1} \sum_{1}^{N} (y_i - \bar{Y})^2$$

$$= \frac{(-1 \cdot 5)^2 + (-0 \cdot 5)^2 + (0 \cdot 5)^2 + (1 \cdot 5)^2}{3}$$

$$= \frac{5 \cdot 00}{3} = 1 \cdot 67$$

எனவே,

$$V(\bar{y}) = \frac{1 \cdot 67}{2} \times \frac{2}{4}$$

$$= \frac{1 \cdot 67}{4} = 0 \cdot 417$$

இரண்டும் சமமாக அமைகின்றன.

பயிற்சி : $E(s^2) = S^2$ என்று சரிபார்க்க.

(ii) ஈடு செய்யப்படும் முறை (With Replacement).

மாதிரி	$\bar{y}$	$\bar{y} - \bar{Y}$	$(\bar{y} - \bar{Y})^2$
1,1	1·0	— 1·5	2·25
1,2	1·5	— 1·0	1.00
1,3	2·0	— 0·5	0·25
1,4	2·5	0	0

2,1	1·5	— 1·0	1·00
2,2	2·0	— 0·5	0·25
2,3	2·5	0	0
2,4	3·0	0·5	0·25
3,1	2·0	— 0·5	0·25
3,2	2·5	0	0
3,3	3·0	0.5	0·25
3,4	3·5	1·0	1·00
4,1	2·5	0	0
4,2	3·0	0·5	0·25
4,3	3·5	1·0	1 00
4,4	4·0	1·5	2·25
			10·00

$$E(\bar{y}) = \frac{40 \cdot 0}{16} = 2 \cdot 5 = \bar{Y}$$

$$V(\bar{y}) = \frac{10 \cdot 00}{16} = 0 \cdot 625$$

மேலும், வரையறைப்படி,

$$V(\bar{y}) = \frac{N-1}{N} \cdot \frac{S^2}{n}$$

$$= \frac{3}{4} \cdot \frac{5}{3} \cdot \frac{1}{2} = \frac{5}{8} = 0 \cdot 625$$

$$\left[S^2 = \frac{5}{3} \right]$$

எல்லா மாதிரிகளிலும் s^2 — மதிப்பிட்டு, அதன்மூலம் $E(s^2) = S^2$ என்று காட்டுக.

எடுத்துக்காட்டு (2):

முழுமைத் தொகுதியானது, மூன்று மாணவர்களைக் கருதி, அவர்தம் கையிருப்புத் தொகைகள் 20, 30, 40 ரூபாய்களானால், இதிலிருந்து எடுக்கப்படும் $n = 2$ அளவில் மாதிரிகளை நிர்ணயித்து, $\bar{y}$-ன் மதிப்பைக் கண்டு, $V(\bar{y})$ என்பதையும் மதிப்பிடுக. (i) $E(\bar{y}) = \bar{Y}$ எனவும்

(ii) $E(s^2) = S^2$ எனவும் நிறுவுக.

மாதிரி எண்	மதிப்புகள் y_1	y_2	$\bar{y}$	$s^2 = \frac{1}{n-1}\sum(y_i-\bar{y})^2$
1	20	30	25	50
2	20	40	30	200
3	30	40	35	50

$\binom{3}{2} = 3$ வழிகளில் ஈருறுப்பு மாதிரிகளைத் தெரிவு செய்யலாம்.

$$\bar{Y} = \frac{20 + 30 + 40}{3} = 30$$

$$E(\bar{y}) = \frac{25 + 30 + 35}{3} = 30$$

$$\therefore E(\bar{y}) = \bar{Y}$$

மேலும் முழுமைத் தொகுதிபற்றிய

$$S^2 = \frac{1}{N-1}\Sigma(y_i - \bar{Y})^2$$

$$= \frac{(-10)^2 + 0^2 + 10^2}{2} = \frac{200}{2} = 100$$

மாதிரி அடிப்படையில்

$$E(s^2) = \frac{50 + 200 + 50}{3} = \frac{300}{3} = 100$$

$\therefore E(s^2) = S^2$ என நிருபணம் ஆகிறது.

$\therefore \bar{y}, s^2$ ஆகியவை, முறையே $\bar{Y}, s^2$ ஆகிய அளவைகளின் பிறழ்ச்சியற்ற மதிப்பீடுகளாகின்றன.

பயிற்சி: $V(\bar{y})$ மதிப்பை இருவழியில் காண்க.

பயிற்சி

1. $N = 600$ என்று உள்ள மாணவர்களினின்று $n = 20$ என்ற அளவில், ராண்டம் மாதிரியினை ஈடு செய்யப்படாத முறையில் தெரிவு செய்க. ராண்டம் எண் பட்டியலைப் பயன் படுத்தும் வெவ்வேறு வழிகளை விவரிக்க.

2. (1. 3, 5, 7, 9) என்று கொடுக்கப்படும் முழுமைத் தொகுதியினின்று $n=2$ என்ற அளவில் ராண்டம் மாதிரிகளை (i) ஈடு செய்யப்படாத முறை, (ii) ஈடு செய்யப்படும் முறை இரண்டின்படியும் தெரிவு செய்க. எல்லாவற்றிற்கும் சராசரிகளையும் மாறுபாட்டு அளவுகளையும் தீர்மானிக்க.

(i) $V(\bar{y}_{ran})$— ன் மதிப்பினை அறிக.

(ii) $E(y_{ran}) = \overline{Y}$ என்று காட்டுக.

(iii) $E(s^2) = S^2$ என்றும் காட்டுக. (ஈடு செய்யப் படாத முறையில்).

3. 6 உறுப்புகளின் மதிப்புக்களை முறையே 0·1, 0·5, 0·8. 1·2, 1·5, 1·9 என்று கொண்ட முழுமைத் தொகுதியினின்று 3 உறுப்புகள் கொண்ட ராண்டம் மாதிரிகள் அனைத்தையும் ஈடு செய்யப்படாத முறையில் தெரிவு செய்து,

(i) மாதிரி சராசரியானது முழுமைத் தொகுதி சராசரியின் பிறழ்ச்சியற்ற மதிப்பீடு என்று காட்டு.

(ii) இச் சராசரியின் மாதிரிமுறை சார்ந்த மாறுபாட்டு அளவினைக் கணக்கிடுக.

(iii) ஒவ்வொரு மாதிரியிலும் s^2 மதிப்பைக் கணித்து $E(s^2) = S^2$ என்று காட்டுக.

(iv) தொகுதி மொத்தமான Y-க்கு ஒரு மதிப்பீட்டினைப் பெறுக. அதன் மாதிரி சார்ந்த மாறுபாட்டினைக் கணக்கிடுக.

4. இயல்நிலைப் பரவலைக் கருதி, தொகுதிச் சராசரி $\mu = 60$ என்றும், திட்டவிலக்க வர்க்கம் $\sigma^2 = 25$ என்றும் உள்ள தொகுதியிலிருந்து $n = 30$ என்ற ராண்டம் மாதிரியினைத் தெரிவு செய்து, அதன் சராசரி மற்றும் மாறுபாட்டு அளவு இவற்றைக் காண்க.

5. 5 மாணவர்களிடம் இருக்கும் புத்தகங்கள் முறையே 30, 50, 60, 70, 80 என்றிருப்பின்,

(i) இம் முழுமைத் தொகுதியின் $\bar{Y}$, σ^2, S^2 இவற்றைக் காண்க.

(ii) இதிலிருந்து எடுக்கப்படும் $n = 2$ அளவினதான மாதிரிகளை (ஈடு செய்யப்படாத சம வாய்ப்பு முறை) கருதி $V(\bar{y})$ மதிப்பை இருவழியில் ஆராய்க

(iii) மேற்கூறிய மாதிரிகளை, ஈடு செய்யப்படும் சம வாய்ப்பு முறையில் எடுத்து $V(\bar{x})$ ன் மதிப்பைக் காண்க.

(iv) இவற்றுக்கு $E(\bar{y}) = \bar{Y}$ என்றும், $E(s^2) = S^2$ என்றும் காட்டுக.

(v) $\hat{Y}$ மற்றும் $\hat{\sigma}_y$ இவற்றைக் காண்க.

6. $N = 50$ என்ற அளவினதான முழுமைத் தொகுதியிலிருந்து $n = 10$ என்ற அளவினதான மாதிரியானது ஈடு செய்யப்படாத ராண்டம் முறையில் தேர்ந்தெடுக்கப்பட்டு,

12, 10, 4, 8, 7, 11, 5, 9, 10 .7

எனும் மதிப்புகளைக் கொண்டிருக்கிறது.

(i) $\bar{y}$-ஐ கணக்கிட்டு, $\hat{Y}$-ஐ கண்டுபிடி.

(ii) $$\sum_{1}^{n} (y_i - \bar{y})^2 = \sum y_i^2 - \frac{(\Sigma y_i)^2}{n}$$

என நிறுவுக.

(iii) s^2 மதிப்பினைக் காண்க.

(iv) $V(\hat{Y})$ என்பதன் மாதிரி மதிப்பீடு காண்க.

7. $N = 250$ என்றமைந்த ஒரு கிராமப்புறக் குடும்பங்களுள், $n = 15$ குடும்பங்களை ஈடு செய்யப்படாத முறைப்படி, சாதாரண ராண்டம் மாதிரியாகத் தெரிவு செய்கையில், அக் குடும்பங்களில் காணப்பட்ட நபர்கள் வருமாறு:

7, 6, 8, 12, 8, 9, 5, 9, 13, 10, 8, 7, 12, 10, 3.

அக் கிராமத்தில் உள்ள மொத்த நபர்கள் எண்ணிக்கையை மதிப்பீடு செய்க. மேலும் இம் மதிப்பீட்டின் தரப்பிழை அளவினையும் கணிக்க.

8. $N = 6$ என்ற முழுமைத் தொகுதி மதிப்புகள் 8, 3, 1, 11, 4, 7 என்றால் இதிலிருந்து பெறக்கூடிய $n = 2$ என்ற அளவினதான எல்லா மாதிரிகளையும் ராண்டம் முறையில் கருதி $\bar{y}$ மதிப்பினைக் கணக்கிட்டு, $E(\bar{y}) = \bar{Y}$ என்று காட்டுக.

மேலும் எல்லா எளிய ராண்டம் மாதிரிகளையும் கருதி, s^2 மதிப்பினைப் பெற்று, $E(s^2) = S^2$ என்று காட்டுக. இதற்கு $n = 3$ என்று கொள்ளவும்.

9. 500 பேர்கள் கொண்ட தொகுதியிலிருந்து 20 பேர்கள் கொண்ட சம வாய்ப்பு மாதிரி எடுக்கப்பட்டு, பெற்ற மதிப்புகள்:

29	38	42	44	51	53	53	54	56	57
60	61	58	59	45	70	65	63	50	45

எனின், முழுமைத் தொகுதி மொத்தத்திற்கு, மாதிரி மதிப்பீடு காண்க. மேலும் அம் மதிப்பீட்டின் மாதிரி சார்ந்த மாறுபாட்டினையும் கணக்கிடுக.

10. ஒரு பெரிய நகரில் 800 பிரிவுகள் உள்ளன. தெருக்களில் நிறுத்தி வைக்கப்படும் கார்களின் எண்ணிக்கையினை ஒரு குறிப்பிட்ட நேரத்திற்குக் கணக்கிட வேண்டியுள்ளது. $n = 25$ என்ற அளவில் பிரிவுகள் தெரிவு செய்யப்படுகின்றன. அவற்றில் காணப்பட்ட கார்கள் விவரம் வருமாறு:

8, 9, 4, 2, 6   6, 3, 0, 8, 4

3, 6, 4, 2, 3   3, 9, 3, 0, 1

4, 2, 3, 6, 1

(*i*) மொத்த கார்களின் எண்ணிக்கையை மதிப்பிடுக

(*ii*) s-ன் மதிப்பு என்ன ?

(*iii*) $\hat{V}(\hat{Y})$ கண்டுபிடித்து அதிலிருந்து $\hat{Y}$-ன் தரப்பிழையினை அறிக.

(*iv*) முழுமைத் தொகுதி மொத்தத்திற்கு 3σ நம்பக எல்லையை வரையறுக்க.

(குறிப்பு: $\hat{Y} \pm 3 . s_{\hat{Y}}$ மதிப்பிடுக.)

11. கீழ்க்கண்ட எண்களிலிருந்து அளவு 4 உள்ள முடிந்த எல்லா சிக்கலற்ற (எளிய) சம வாய்ப்பு கூறுகளை ஈடுசெய்யப் படாத முறையில்) எடுக்கவும். தொகுதியின் சராசரிக்குக் கூறின் சராசரி பாரபட்சமற்ற மதிப்பீடு என்பதையும், தொகுதியின் சராசரி வர்க்கத்திற்கு கூறின் சராசரி வர்க்கம் பாரபட்சமற்ற மதிப்பீடு என்பதையும் சரிபார்க்கவும்.

10·4, 9·8, 11·2, 10·6, 8·0

[B.Sc. Sep., 75 (Anc)]

12. ஒரு முழுமைத் தொகுதியின் இயல்பானது (x), $x = \alpha + \beta\,\theta$ என்ற நேர்கோட்டுப் போக்கில் அமைகிறது. α. β என்பவை மாறிலிகள் (Constants) $\theta = 0, 1, 2 \ldots\ldots (N-1)$. முழுமைத் தொகுதி சராசரியான $\bar{x}_N$—க்கு ஒரு மதிப்பீடு பெற வேண்டியுள்ளது. எத்தகைய ஒரு மாதிரி முறையினை இதற்குப் பயன் படுத்துவது என்று விளக்குக. காரணங்களையும் கூறுக.

(B.Sc., Sep., 75.)

13. சிக்கலற்ற சமவாய்ப்புக்கூறு எடுத்தல் என்பதை விளக்குக. 250 வீடுகளைக் கொண்ட ஒரு கிராமத்திலிருந்து ஈடு செய்யப் படாமல் 15 வீடுகளைக் கொண்ட ஒரு சிக்கல் இல்லாத சம வாய்ப்புக் கூறு எடுக்கப்பட்டிருக்கிறது. கூறில் இருந்த வீடுகளில் இருந்த ஆட்களின் எண்ணிக்கை

7, 6, 8, 12, 8, 9, 5, 9, 13, 10, 8, 7, 12, 10. 3. கிராமத்தின் ஜனத் தொகையினை மதிப்பிட்டு, மதிப்பீட்டின் திட்டப்பிழையினைக் காண்க.

கூறிய 8 வீடுகளில் சைக்கிள் ஒவ்வொன்று காணப்பட்டது. கிராமத்தில் மொத்தம் எவ்வளவு வீடுகளில் சைக்கிள் உள்ளது என்பதை மதிப்பிடவும்.

(B. Sc., (Anc). Ap-74)

வீதம் மற்றும் சதவீதம்பற்றிய மாதிரிமுறை (Sampling for (Proportion & Percentage)

முழுமைத் தொகுதியானது, பல நேரங்களில் ஏதேனும் ஒரு பண்பின் (attribute) அடிப்படையில் இரு பிரிவுகளாகப் பகுக்கப் பட்டதாக இருக்கும். வரையறுக்கப்பட்ட பிரிவுகளில், ஏதேனும் இயல்பானது காணப்படுவதற்கேற்ப பாகுபாடு அமையும். உதாரணமாக, ஒரு குறிப்பிட்டபரப்பில் உள்ள மொத்த மக்களில்,

எழுதப் படிக்கத் தெரிந்தவர்களின் சதவீதமானது கொடுக்கப் படலாம். மக்கட்தொகை கணிப்பு ஆய்வுகளில், இம் மாதிரிப் பாகுபாடு அதிக அளவில் காணப்படுகிறது. இத்தகைய இரு வழிப் பாகுபாட்டில், மாதிரிமுறையில் சிறந்த இடம்பெறுகின்றன.

அதாவது முழுமைத் தொகுதியின் கூட்டு மொத்த அளவு வீதஅளவு (Proportion), அல்லது சதவீதமதிப்பீடு (Percentage) போன்றவை அறியப்பட நேரிடலாம். தொகுதியானது இரு கூறுகப் பிரிக்கப்பட்டு, வீத ஆய்வு நடைபெறுகிறது. புள்ளி விவர ஆய்வு முறைகள் ஒவ்வொன்றிலும், வீதம்பற்றிய மதிப்பீடுகள் தனி இடத்தை வகிக்கின்றன.

இதன்படி, முழுமைத் தொகுதியின் ஒவ்வொரு கூறும் (unit) அல்லது அலகும், கருதப்படும் c, c' என்ற இருபிரிவுகளில் ஏதேனும் ஒன்றில் இடம் பெறுவதாகக் கொள்கிறேம்.

{ மாறியிணைப் பயன்படுத்தி, $y_i = 0$ என்றால், கருதப்பட்ட i-ஆவது கூறுனது ஒரு குறிப்பிட்ட பிரிவில் இடம் பெறவில்லை என்றும், $y_i = 1$ எனில், அக்கூறு குறிப்பிட்ட பிரிவில் இடம் பெறுகிறதெனவும் கொள்வது கணித முறை ஆய்வுக்குப் பெரிதும் உதவுகிறது.} C என்ற பிரிவில், மொத்தமாக, A அளவு முழுமைத் தொகுதிக் கூறுகள் இடம் பெற்றிருப்பின் முழுமைத் தொகுதி வீத அளவு $\frac{A}{N}$ என்றும் (N : முழுமைத் தொகுதியின் அளவு) அதே C பிரிவில், a-மாதிரிக் கூறுகள் அமைந் திருப்பின், மாதிரி வீத அளவு $\frac{a}{n}$ என்றும். N அளவு தொகுதியி லிருந்து n-அளவுள்ள மாதிரிக்கூறு தெரிவு செய்யப்படுகிறது.) வரையறுக்கலாம்.

விகிதசம மதிப்பீடு (Estimation of Proportion)

ஒரு குறிப்பிட்ட பிரிவுக்குள் அடங்கும் (class) முழுமைத் தொகுதிக் கூறுகளின் விகிதசமஅளவான P என்பதைநிர்ணயித்தல் சில ஆய்வுகளில் தேவைப்படுகிறது. முழுமைத் தொகுதியானது ஏதேனும் ஒரு பண்பின் (attribute) அடிப்படையில் இரு பிரிவு களாகப் பகுக்கப்பட்டு, ஒரு கருதப்பட்ட பிரிவான C என்பதில் N' கூறுகள் இருப்பதாகக் கொண்டால், $P = \frac{N'}{N}$ என்ற விகிதசம அளவினை முழுமைத் தொகுதிக்கு வரையறுக்கிறேம்.

உதாரணமாக, ஒரு இடத்தில் உள்ள மக்களுள், வேலை யற்றோர் (unemployed) வீத அளவினை (Percentage) அல்லது ஒரு நோயினால் பாதிக்கப்பட்டோர் வீதத்தை மதிப்பிட நேரிடலாம். இது போன்ற நிலைகளில், முழுமைத் தொகுதி இயல்புகள்

$$\bar{Y} = \frac{N'}{N} = P \quad (\because Y = N')$$

எனவும்,

$$\sigma^2 = \frac{1}{N}\sum_{1}^{N} Yi^2 - Y^2$$

$$= P - P^2 = PQ, \; (Q = 1 - P)$$

எனவும் பெறப்படும்.

{ i — எனும் உறுப்பின் மதிப்பு (Y_i) கருதப்பட்ட பிரிவில் இருப்பின், $Y_i = 1$ எனவும், இல்லையெனில் $Y_i = 0$ எனவும் கொள்க. }

ஈடு செய்யப்பட்ட சமவாய்ப்பு மாதிரி முறையில், மாதிரி விகித அளவு (Proportion) p என்பது

$$\bar{y} = p = \frac{r}{n}$$

எனப்படும். r என்பது, குறிப்பிட்ட பிரிவில் உள்ள மாதிரிக் கூறுகள் (units) ஆகும். $E(p) = P$ ஆகும்.

மேலும் $V(p) = \frac{PQ}{n}$ என்று பெறப்படும்.

முழுமைத் தொகுதி மாறுபாட்டுக் கெழு (coefficient of variation)

$$C = \sqrt{\frac{PQ}{P}} = \sqrt{\frac{Q}{P}} \text{ ஆகும்.}$$

எனவே $C_p = \frac{1}{\sqrt{n}}. \sqrt{\frac{Q}{P}}$ ஆகும். (தொடர்புடைய மாதிரிப் பிழை அளவு).

σ^2-ன் பிறழ்ச்சியற்ற மதிப்பீடு s^2 என்பது

$$s^2 = \frac{1}{n-1}(np - np^2)$$

$$= \frac{npq}{n-1} \quad (q = 1 - p)$$

என்று பெறப்படும். இது PQ-ன் பிறழ்ச்சியற்ற மதிப்பீடு.

$V(p)$ உடைய பிறழ்ச்சியற்ற மதிப்பீடு

$$v(p) = \frac{pq}{n-1}$$

என்று பெறப்படும்.

ஈடு செய்யப்படாத முறையில்,

$$V(p) = \frac{N-n}{(N-1)} \cdot \frac{PQ}{n} \text{ என்றும்}$$

$$C_p = \sqrt{\frac{N-n}{N-1}} \cdot \frac{1}{\sqrt{n}} \cdot \sqrt{\frac{Q}{P}} \text{ என்றும்}$$

பெறப்படும்.

இங்கு, $S^2 = \dfrac{N\sigma^2}{N-1}$ என்பதால்,

$S^2 = \dfrac{N}{N-1} \cdot PQ$ என்று அடையப்பெற்று, S^2-ன் பிறழ்ச்சியற்ற மாதிரி மதிப்பீடு (Sample Estimate),

$$s^2 = \frac{n}{n-1} \cdot pq$$

என்று கணக்கிடப்படும்.

மேலும், $V(p)$ என்ற மாறுபாட்டின் பிறழ்ச்சியற்ற மாதிரி மதிப்பீடு

$$V(p): \ v(p) = \frac{pq}{n-1} \cdot \frac{N-n}{N} = \frac{pq}{n-1}(1-f)$$

என்று கணக்கிடப்படும்.

மேற்கண்ட முடிவுகளை, பின்வரும் எடுத்துக்காட்டில் விளக்கலாம்:

எடுத்துக்காட்டு 1 :

கருதப்பட்ட 5 நபர்களுள் 3 பேர் படித்தவர்கள் ஆவர். $n = 2$ அளவுள்ள ராண்டம் மாதிரிகளைத் தேர்ந்தெடுத்து, (திரும்ப வைக்காத முறைப்படி), சில முடிவுகளைப் பெறுவோம்.

வழிமுறை :

$N = 5$; $n = 2$. எனவே மொத்தம் $\binom{5}{2} = 10$ மாதிரிகள் எடுக்கப்படலாம்.

$Y_i = 1$ என்பது i எனும் கூறு கருதப்பட்ட பிரிவிலிருப்பின்.
$= 0$ .... இல்லாவிடின்.

இதன்படி, $Y = 1, 1, 1, 0, 0$ என்ற மதிப்புகள் அமையும்.

மாதிரி	y-ன் மதிப்புகள்	Σy	$p = \frac{\Sigma y}{n}$	pq
1	1, 1	2	1·0	0·0
2	1, 1	2	1·0	0·0
3	1, 0	1	0·5	0·25
4	1, 0	1	0·5	0·25
5	1, 1	2	1·0	0·0
6	1, 0	1	0·5	0·25
7	1, 0	1	0·5	0·25
8	1, 0	1	0·5	0·25
9	1, 0	1	0·5	0·25
10	0, 0	0	0·0	0·0

(1) $P = \frac{N'}{N} = \frac{3}{5} = 0{\cdot}6 \; (N' = 3)$.

இது படித்தவர்கள் வீதம்.

மாதிரி அடிப்படையில்,

$E(p) = p$-ன் சராசரி மதிப்பு

$$= \frac{\Sigma p}{10} = \frac{6}{10} = \frac{3}{5} = 0{\cdot}6.$$

எனவே p-ஆனது, P-ன் பிறழ்ச்சியற்ற மதிப்பீடு.

(2) Y_i-ன் மாறுபாட்டு அளவு (Variance)

$$S^2 = \frac{N}{N-1} \cdot PQ$$

$$= \frac{5}{4} \cdot (0{\cdot}6) \cdot (0{\cdot}4) \qquad Q = 1 - P = 1 - 0{\cdot}6 = 0\ 4$$

$$= 0{\cdot}3$$

S^2-ன் மாதிரி மதிப்பீடு (Sample estimate of S^2)

$$s^2 = \frac{n}{n-1} \cdot pq$$

$$= \frac{2}{1} \cdot pq$$ (p-ன் வெவ்வேறு மாதிரி மதிப்புகட்கு கணக்கிடப்படும்.)

$$E(s^2) = \frac{2}{1} \left[\frac{0 + 0 + 0{\cdot}25 + \ldots\ldots + 0{\cdot}0}{10}\right]$$

$$= \frac{2 \times (1{\cdot}50)}{10} = 0{\cdot}3 = S^2$$

எனவே s^2 என்பது S^2-ன் பிறழ்ச்சியற்ற மதிப்பீடாகும்.

(3) p மதிப்பின் மாதிரிமுறைசார் மாறுபாடு (Sampling Variance)

$$V(p) = \frac{PQ}{n} \cdot \frac{N-n}{N-1}$$

$$= \frac{(0.6)(0\ 4)}{2} \cdot \frac{(5-2)}{(5-1)}$$

$$= 0{\cdot}09$$

இதன் மாதிரி மதிப்பீடு (Sample Estimate)

$$\hat{V}(p) = v(p) = \frac{pq}{n-1} \cdot \frac{N-n}{N}$$ (வெவ்வேறு p-மதிப்புகட்கு)

$$\therefore \quad E[v(p)] = \frac{1}{n-1} \cdot \frac{N-n}{N} \cdot E(pq)$$

$$= \frac{1}{(2-1)} \frac{(5-2)}{5} \left[\frac{0 + 0 + 0{\cdot}25 + \ldots\ldots\ldots + 0{\cdot}0}{10}\right]$$

$$= \frac{3}{5} \times (0{\cdot}15)$$

$$= 0{\cdot}09 = V(p)$$

∴ இதன் பிறழ்ச்சியின்மையும் புலப்படும்.

எடுத்துக்காட்டு 2 :

$N = 1000$. $n = 50$ அளவுள்ள மாதிரி எடுக்கப்பட்டால், அதில் 20 உறுப்புகள் ஒரு பிரிவில் அடங்கும். P-யின் மதிப்பை மதிப்பிடுக.

நியதிப்படி, P என்ற விகிதசமத்தின் மதிப்பீடு, மாதிரி மதிப்பான p ஆகும். ($\because \; E(p) = P$)

$$\therefore \quad p = \frac{20}{50} = 0{\cdot}4 \text{ ஆகும்.}$$

இதுவே P-ன் பிறழ்ச்சியற்ற மதிப்பீடு.

பயிற்சி

1. 5 மாணவர்களில், 3 பேர் புகைபிடிப்பவர்களாக அமைகின்றனர். Y-ன் மதிப்புகளை (1, 1, 1, 0, 0) என்று கருதுக. ['1' எனில் புகைபிடிப்பவராகவும், '0' எனில் புகை பிடிக்காதவராகவும் அறியப்படும்]. $n = 2$ என்ற அளவு மாதிரிக் கூறுகளை, ஈடு செய்யப்படாத முறையில் தெரிவு செய்து முழுமைத் தொகுதி விகிதமான P யின் மதிப்பீடுகளைப் பெறுக.

இதில், மாதிரிகளை ஈடு செய்யப்படும் முறையில் தேர்ந்தெடுக்கையில், $E(p) = P$ என்று காட்டுக.

$$\left\{ \text{குறிப்பு :} \quad p = \frac{\Sigma y_i}{n} \text{ என்று பயன்படுத்துக.} \right\}$$

2. மேற்கண்ட பயிற்சி (1)-ல், $n = 2$ என்ற அளவில் எடுக்கப்பட்ட மாதிரிகட்கு $\hat{V}(p)$ யின் மதிப்புகளைப் பெறுக. மேலும் $\hat{V}(p)$ மதிப்பீடு பிறழ்ச்சியற்றது என்று நிறுவுக.

{ அதாவது $E[\hat{V}(p)] = V(p)$ என்று காட்டுக.}

3. முழுமைத் தொகுதியில் $N = 6$; $A = 4$; $A' = 2$ என்று தரப்படுகிறது. $n = 3$ என்ற அளவில் எடுக்கப்படும்

எல்லா ராண்டம் மாதிரிகளையும் கருதி, a-ன் மதிப்பை கணிக்க. மேலும், P என்ற] விகிதத்தின் மாதிரிச் சராசரி மற்றும் அம் மதிப்பீட்டின் மாறுபாடு ஆகியவற்றைக் கணக்கிடுக. மேலும்,

$$V(p) \text{ யின் பிறழ்ச்சியற்ற மதிப்பீடு } \frac{N-n}{N} \cdot \frac{1}{N-n} \cdot pq$$

எனக் காட்டுக.

{ A, A' → தொகுதி உட்பிரிவு அளவுகள் ; a - மாதிரி உட்பிரிவு அளவு }.

சம வாய்ப்பு மாதிரி முறையில் மாதிரி அளவும் திட்பமும் (Sample Size and Precison in Random Sample Method)

பொதுவாக, திட்பமானது $d = z.\sigma$ என்று ஒரு மாதிரி மதிப்பீட்டு அளவைக்கு (estimator) வரையறுக்கப்பம். σ என்பது கருதப்பட்ட மாதிரி அளவையின் தரப்பிழை; z என்பது நம்பகத் தன்மைபற்றிய காரணி. ஈடு செய்யப்படும் (with replacement) மாதிரி முறையில், $\bar{y}$ என்ற மாதிரிச் சராசரி அளவையின் திட்பமானது

$$d = z \frac{s}{\sqrt{n}}$$

என்று பெறப்படும். s என்பது,

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (y_i - \bar{y})^2$$

என்று வரையறுக்கப்பட்டு, $\frac{s^2}{n}$ ஆனது $\bar{y}$-ன் முழுமைத் தொகுதி மாறுபாட்டின் (variance) பிறழ்ச்சியற்ற மதிப்பீடாகும். இதிலிருந்து, மாதிரி அளவினை,

$$n = z. \frac{(z. s)^2}{d^2} \qquad \dots (1)$$

என்று பெறுகிறோம்.

ஈடு செய்யப்படாத (without replacement) மாதிரி முறையில் வரம்புடைய முழுமைத் தொகுதி திருத்தம் (Finite Population Correction) பயன்படுத்தப்பட்டு,

$$d = z. \frac{s}{\sqrt{n}} . \sqrt{\frac{N-n}{N}}$$

என்றுக இதிலிருந்து

$$n = \frac{N. (zs)^2}{Nd^2 + (zs)^2} \quad \ldots\ldots (2)$$

என்று மதிப்பிடப்படும்.

உதாரணமாக மிக அதிக அளவில் (very large) அமைந்த தொகுதியின் மாறுபாடானது $s^2 = 600$ என்று மதிப்பிடுவதாகக் (estimated) கொள்க. இத் தொகுதியின் சராசரியினை ± 5 என்ற திட்ட வரையறையில், 99% நம்பகத் தன்மையில் மதிப்பிட, எவ்வளவு பெரிய அளவினதான மாதிரியினைக் கருதவேண்டும் ?

இங்கு, $d = \pm 5$; $s = \sqrt{600}$;

$z = 3$ (99% அளவில்).

$$\therefore \quad n = \frac{9 \times 600}{25} = 216$$ என மதிப்பிடப்படும்.

(முடிவு (1)ன் மூலம்.)

$\therefore \bar{Y}$ என்ற முழுமைத் தொகுதி சாரசரியின் நம்பக எல்லை.

$$y - z. \frac{s}{\sqrt{n}} < \bar{Y} < y + z. \frac{s}{\sqrt{n}}$$

என்பதாகும். அதில், $z \frac{s}{\sqrt{n}} = \frac{3 \times \sqrt{600}}{\sqrt{216}} = \sqrt{25} = 5$

அதாவது மதிப்பிட்ட சராசரியானது, $\bar{Y}$ — லிருந்து ± 5 அளவில் வேறுபட்டு இருக்கும்.

மாறுக, $N = 2000$ எனக் கருதுவோம். இதில் *fpc* யினை அதாவது, வரம்புடைய முழுமைத் தொகுதி திருத்தத்தைப் பயன் படுத்தி (ஏனெனில் எண்ணிக்கை வரம்புடைய முழுமைத்தொகுதி கருதப்படுகிறது). தேவையான மாதிரியின் அளவினை

$$n = \frac{200 \times 9 \times 600}{(2000 \times 25) + (9 \times 600)}$$

$$= 195$$

என்று மதிப்பிடுகிறோம். [முடிவு (2) பயன்பட்டது.]

பயிற்சி

(1) 2000 குடும்பங்கள் கொண்ட ஒரு தொகுதியிலிருந்து, ஈடுசெய்யப்படாத முறையில் ஒரு மாதிரியைத் தெரிவுசெய்து உணவுக்காக ஆகும் ஒரு வாரச் சராசரிச் செலவினை மதிப்பிட, முழுதும் நிகழத்தக்க நிலையில் (almost certain), ± \$ 4 என்ற திட்பத்துடன் கூடிய மாதிரியின் அளவு என்ன? $s = \$\ 12$ என்று கொடுக்கப்படுகிறது.

(2) ஒரு வட்டத்தில் 500 பண்ணைகள் உள்ளன. ஒவ்வொரு பண்ணையிலும் சராசரியாக உள்ள மாடுகள் கணக்கிட வேண்டியுள்ளது. எடுக்கப்பட்ட 30 ராண்டம் மாதிரிப் பண்ணைகளில் ($n = 30$) கண்ட விவரங்கள் :

23, 14, 38, 11, 7	31, 9, 16, 12, 25
11, 15, 36, 24, 10	16, 26, 12, 25, 28
34, 25, 11, 7, 9	33, 25, 9, 4, 19

(a) $\bar{y}$-ஐ மதிப்பிட்டு அறிக.

(b) s—ன் மதிப்பைக் கண்டு அதன்வழி $S\bar{y}$ மதிப்பு காண்க.

(c) இந்த s-ஐப் பயன்படுத்தி, ± 4 அளவில் திட்பமுடைய, $z = 3$க்கு ஏற்ற, மாதிரி அளவினை நிர்ணயிக்க.

3. ஈடு செய்யப்படும் முறையில் மாதிரி எடுக்கப்படுகையில் n_0 என்பதை மாதிரி அளவாகவும், ஈடு செய்யப்படாத முறையில் n என்றும் கொண்டு.

$$n = \frac{n_0}{1 + \frac{n_0}{N}}$$ என்று காட்டுக.

[குறிப்பு : $n_0 = \frac{z^2\ s^2}{d^2}$ என்று கொண்டு,

$$n = \frac{N\,(zs)^2}{Nd^2 + (zs)^2}$$ என்பதில் தக்கபடி பிரதியிடவும்.]

4. பயிற்சி (1)-ல், n_0 மதிப்பை அறிந்து, (3)-ல் காட்டப்படும் சூத்திரத்தின் மூலம் n-மதிப்பினை கணக்கிடுக.

5. (*i*) தொகுதி மொத்தத்தினை (Y) மதிப்பீடு செய்கையில், அதன் திட்பத்திற்கேற்ப, மாதிரி அளவினை

$n = \dfrac{n_0}{1 + \dfrac{n_0}{N}}$ என்று தீர்மானிப்பதாக காட்டுக.

$$\left(n_0 = \frac{N^2 \cdot (zs)^2}{d^2} \right).$$

(*ii*) மாறுபாட்டுக் கெழுவான c^2-ஐப் பயன்படுத்தி (Coefficient of Variation).

$$d_0{}^2 = z^2 \cdot C_{\hat{Y}}{}^2 \quad \left(d_0 = \frac{\hat{Y} - Y}{Y} \right)$$

என்று கொடுக்கப் படுகையில்,

$$n = \frac{N\,(zc)^2}{Nd_0{}^2 + (zc)^2}$$

என்று வருவதை நிரூபிக்கவும்.

6. $N = 1000$; திட்பமானது, கூட்டு மொத்தத்தினின்று $\pm$ 10%-ல் அமைவதாக வேண்டும். முன் கூட்டி நடைபெறும் ஆய்வில் $C = 0{\cdot}3$ என்று தீர்மானிக்கப்பட்டால், $z = 3$ என்ற நம்பக அளவிற்கேற்றபடி, மாதிரி அளவு எவ்வளவு பெரியதாக (சமவாய்ப்பு முறையில்) இருக்க வேண்டும் ?

7. முழுமைத் தொகுதியின் அளவு 2000 என்றும், அதன் மாறுபாட்டு அளவின் மாதிரி மதிப்பீடு $s^2 = 144$ என்றும் தரப்படின், $\pm$ 4 என்ற திட்ப அளவில், 99% நம்பகத் தன்மை யுடன் n-ன் மதிப்பைக் காண்க.

விதம் பற்றிய நம்பக எல்லை, திட்பம் மற்றும் மாதிரி அளவு (Confidence Interval, Precision & Sample Size for Proportion)

நம்பக எல்லை

p-மதிப்பின் திட்டப்பிழை (Standard error) கணிக்கப்பட்ட பின், முழுமைத் தொகுதி மதிப்பு P-க்கு, ஒரு குறிக்கப்பட்ட நம்பகத் தன்மையில், நம்பக எல்லையினை

$$p - z . s_p < P < p + z \cdot s_p$$

என்று வரையறுக்கலாம். p-ன் மாதிரிக் கூற்றுப் பரவல் (Sampling distribution), இயல்நிலையாக (normal) அமைந்திருப்பதாகக் கொள்ளப்படுகிறது. இது பற்றிய விளக்கத்தினை, மாதிரி அளவு தீர்மானம் பற்றி 'கொக்ரான்' (Cochran) புத்தகத்தில் (Sampling Techniques) காணலாம்.

மாதிரி அளவு மற்றும் திட்ப அளவு :

முழுமைத் தொகுதி வீத அளவான P-யினை மதிப்பிடும் p-மதிப்பீட்டின் திட்ப அளவு,

$$s_p^2 = \frac{pq}{n-1} \cdot \frac{N-n}{N}$$

என்ற தரப்பிழை வர்க்கத்தினைப் பயன்படுத்தி பெறப்படும் $(n-1)$ என்பதை n என்று கொள்வது, கணக்கீட்டை எளிதாக்கும்.

∴ ஈடு செய்யப்படாத (without replacement) முறையில்,

$$d^2 = z^2 \cdot \frac{pq}{n} \cdot \frac{N-n}{N}$$ என்றும்,

திரும்ப வைக்கும் முறையில்,

$$d^2 = z^2 \cdot \frac{pq}{n}$$ எனவும்

திட்ப அளவு கணக்கிடப்படும். z என்பது கொடுக்கப்பட்ட 'α' அளவிற்கேற்ப அமையும் நம்பகத் தன்மை காரணி.

இவற்றுக்கு ஏற்ற வகையில் முடிவு செய்யப்படும் மாதிரி அளவு n· ஆனது

$$n = \frac{N \cdot z^2 \cdot pq}{Nd^2 + z^2 \cdot pq}$$ (ஈடு செய்யப்படாத முறையின் படி)

$$n = \frac{z^2 \cdot pq}{d^2}$$ (ஈடு செய்யப்பட்ட முறையின் படி)

என்று கணக்கிடப் பெறும்.

எடுத்துக்காட்டு 1 :

ஒரு ஆய்வில், மாதிரியாகக் கருதப்பட்ட 100 குடும்பங்களில், ஒரு பிரச்சினைக்கு ஆதரவாக 60 குடும்பங்களும், எதிர்த்து ஏனையோரும் அமையும் நிலை அறியப்படுகிறது. முழுமை தொகுதி வீத அளவிற்கு (proportion), அதாவது P-க்கு, 95% நம்பக தன்மையில், நம்பக இடைவெளியினை (Confidence interval) தீர்மானிக்க $N = 5000$ என்று கொள்க.

கொடுக்கப்பட்டவை

$$a = 60$$

$$n = 100$$

$$\therefore \quad p = \frac{a}{n} = \frac{60}{100} = \frac{3}{5} = 0{\cdot}6$$

எனவே $q = 0{\cdot}2$

p-யின் தரப்பிழை

$$s_p^2 = \frac{pq}{n-1} \cdot \frac{N-n}{N}$$

$$= \frac{(0{\cdot}6)\,(0{\cdot}4)}{99} \cdot \frac{4900}{5000}$$

$$= 0{\cdot}0024$$

$$\therefore \quad s_p = \sqrt{0{\cdot}0024} = 0{\cdot}049$$

எனவே, $z = 2$ என்று கொண்டு ($\because$ 95% நம்பக தன்மை) நம்பக எல்லை

$$p - z \cdot s_p < P < p + z \cdot s_p$$

அதாவது

$$p - 2s_p < P < p + 2s_p$$

அதாவது

$$0{\cdot}8 - (2 \times 0{\cdot}049) < P < 0{\cdot}8 + (2 \times 0{\cdot}049)$$

அதாவது

$$0{\cdot}702 < P < 0{\cdot}898$$

என்று பெறப்படுகிறது.

எடுத்துக்காட்டு 2 :

$N = 4000$ என உள்ள தொகுதியினின்று, சமவாய்ப்பு முறையில் தேர்ந்தெடுக்கப்பட்ட 50 மாணவர்களுள் 30 பேர் மிதி வண்டி வைத்திருப்பதாக அறியப்படுகிறது. திட்ப அளவினை $\pm 5\%$ ஆகக் கருதி, P-யின் மதிப்பீட்டினைப் பெற 99% நம்பகத் தன்மையில், தேவைப்படும் மாதிரி அளவு என்ன ?

n-மதிப்பைக் கணக்கிட வேண்டும்.

$$P = \frac{30}{50} = 0{\cdot}6 \;;\quad q = 0{\cdot}4 \;;\qquad d^2 = (0{\cdot}05)^2$$
$$= 0{\cdot}0025.$$

வரையறைப்படி, $z = 3.$

$$n = \frac{N\, z^2 . pq}{Nd^2 + z^2 . pq}$$
$$= \frac{4000 \times (3)^2\ (0{\cdot}6)\,(0{\cdot}4)}{(4000 \times (0{\cdot}0025) + (9 \times 0{\cdot}24)}$$

$\therefore\ n = 716$ என்று அறியப்படுகிறது.

எனவே அதிகப்படியாக 666 மாணவர்களைத் தேர்வு செய்தல் வேண்டும்.

பயிற்சி

1. $N = 5000$ என்று மாணவர்கள் உள்ள தொகுதியினின்று $n = 100$ எனும் அளவு சமவாய்ப்பு மாதிரி தெரிவு செய்யப்படு கிறது. 40 பேரிடம் சொந்த கார் இருப்பதாகக் கணக்கிடப்படு கிறது. $\pm 5\%$ திட்பத்துடன், 95% நம்பக எல்லையில் P-யின் மதிப்பீட்டைப் பெறத் தேவையான மாதிரி அளவினை தீர்மானிக்க.

2. ஒரு கருத்தறிவுப் பகுப்பாய்வில் கருதப்பட்ட மொத்தம் 1000 நபர்களுள், மாதிரியாக எடுக்கப்பட்ட 100 பேரில் 60 பேர் பிரச்னைக்கு ஆதரவாகவும், 30 பேர் எதிர்த்தும், 10 பேர் எக் கருத்தும் கூறுமலும் இருக்கின்றனர். ஆதரவு தெரிவிப்பவர் களின் விகிதத்திற்கு (P), முழுமைத்தொகுதி மதிப்பிற்கு 95% நம்பகத் தன்மையில், நம்பக எல்லையினை நிர்ணயிக்க.

3. 2500 கல்லூரி மாணவர்களிடம், புகை பிடிப்போர் விகிதத்தினை மதிப்பிட வேண்டும். திட்ப அடவு $\pm 8\%$ என்று கருதி, 95% நம்பகத் தன்மையில், தேவைப்படும் மாதிரி அளவு என்ன ?

4. $N = 3000$. திட்ப அளவு $\pm 0{\cdot}06$ என்றும், நம்பகத் தன்மை 95% என்றும் கருதப்படின், மாதிரி அளவு எவ்வளவு பெரிதாக இருத்தல் வேண்டும். 99% நம்பகத் தன்மையுடன் கூடிய நம்பக எல்லையினை $n = 50$; $a = 20$ என்று எடுத்துக் கொண்டு தீர்மானிக்கவும் $\left(p = \frac{a}{n} \right.$ எனக் கொள்க. $\left. \right)$

5. 100 உறுப்புகள் கொண்ட ஒரு சமவாய்ப்பு மாதிரியில், ஒரு பிரிவில் (C எனக் கொள்க) 37 உறுப்புகள் இடம் பெறுகின்றன. முழுமைத் தொகுதியில் 500 உறுப்புகள் இருப்பதாகக் கொண்டு, முழுமைத் தொகுதியின் வீத அளவுக்கும் (proportion) C-பிரிவில் உள்ள மொத்த உறுப்புகட்கும், 95% நம்பக இடைவெளியினைத் தீர்மானம் செய்க.

6. 2000 கல்லூரிகளினின்று, எளிய சமவாய்ப்பு முறையில் தெரிவு செய்யப்பட்ட 200 கல்லூரிகளில் 57 ஒரு கருத்தினுக்கு எதிராகவும், 120 ஆதரவாகவும், 23-ல் எந்த விவரமும் இல்லாமலும் இருப்பதாக அறியப்படுகிறது. கருத்துக்கு ஆதரவாக முழுமைத் தொகுதியில் இடம்பெறும் மொத்தக் கல்லூரிகள் எண்ணிக்கைக்கு 95% நம்பகத்தன்மைடன் கூடிய நம்பக இடைவெளியினை நிர்ணயிக்க.

மாறுபாட்டுக் கெழுவின்மூலம் மாதிரி அளவு கணக்கீடு (Calculation of Sample Size from Coefficient of Variation)

மாறுபாட்டுக் கெழுக்கள் (C.V.)

$$C(\bar{y}) = \frac{C}{\sqrt{n}} \quad \text{(ஈடு செய்யப்படும் மாதிரிமுறையின்படி)}$$

$$C(\bar{y}) = \sqrt{\frac{N-n}{N-1}}\;\frac{C}{\sqrt{n}} \quad \text{(ஈடு செய்யப்படாத முறையின்படி)}$$

என்று இருவகையில் அமைவதை எளிய சமவாய்ப்பு முறையில் காணலாம். இவற்றில் $C = \frac{\sigma}{\bar{Y}}$ (முழுமைத் தொகுதியின் மாறுபாட்டுக் கெழுவாகும்.) இது கொடுக்கப்படுகையில், n—அளவு மதிப்பிடப்படலாம். இதற்கு அனுமதிக்கக்கூடிய பிழையின் எல்லையானது தரப்பட்டிருத்தல் அவசியம். மதிப்பீட்டினது, தொடர்புபடுத்தப்பட்ட திட்டப் பிழை (Relative Standard Error) அளவு 'e' என்று கொண்டு,

$$n_0 = \frac{c^2}{e^2} \quad \text{(திரும்ப வைக்கும் முறை)}$$

என்ற மதிப்பீடு பெறப்படும். இதன்மூலம், திரும்ப வைக்காத முறையில்

$$= \frac{NC^2}{(N-1)\, e^2 + c^2} = \frac{N.n_0}{N + n_0 - 1}$$

என்ற மாதிரி அளவின் மதிப்பீடு பெறப்படும். N மிகப் பெரியதாக இருப்பின், n_0, n இரண்டு மதிப்புக்களும் ஏறக்குறைய சம அளவினதாக அமைகின்றன.

இந்த வரையறைகளில் பயன்படுத்தப்பட்ட 'e' என்னும் மதிப்பானது,

$$P_r \left[\frac{|\bar{y}-\bar{Y}|}{S.E\,(\bar{y})} \leq z_{\alpha} \right] = (1-\alpha)$$

என்ற நிகழ்தகவின் அடிப்படையில் நிர்ணயிக்கப்படுவது.

$$\text{அல்லது } P_r \left[\left| \frac{\bar{y}-\bar{Y}}{\bar{\bar{y}}} \right| \leq z_{\alpha}\, C\,(\bar{y}) \right] = 1-\alpha)$$

என்பதற்கு ஏற்ற $c\,(y)$ மதிப்பு (அதாவது e)

$$e = \frac{d}{z_{\alpha}} \text{ என்பதாகும்.}$$

(d = திட்ப அளவு — கொடுக்கப்படுவது.

எனவே, α மற்றும் d மதிப்புகள் அறியப்பட்ட நிலையில், e-மதிப்பினைப் பெற இயலும். c-யின் மதிப்புடன் இணைக்க, n-மதிப்பீடு கிடைக்கும். நம்பக எல்லையின் அளவு (Confidence interval). $2d$ என்று அமைகிறது.

c-யின் மதிப்பு அறியப்படாத நிலையில், $e\,(\bar{y})$-ன் மாதிரி மதிப்பீட்டினைப் பயன்படுத்தி முடிவுகள் பெறப்படும். அந் நிலையில்,

$$P_r \{ |\bar{y} - \bar{Y}| \leq z_{\alpha} \,.\, s\,(\bar{y}) \} = 1 - \alpha$$

என்ற நிகழ்தகவு (probability) வரையறையானது, கொடுக்கப் படும் மாதிரி அளவிற்குப் பெறப்படும்.

நம்பக எல்லையின் எதிர்பார்க்கும் நீளமானது, $E\,(L)$ என்று குறிக்கப்பட்டு, மாதிரி அளவினையும் மற்றும் முழுமைத்

தொகுதியின் திட்ட விலக்கத்தினையும் (Standard deviation) பொறுத்து அமைந்திருக்கும்.

நம்பக எல்லையின் நீளமானது

$L = 2 \cdot z_{\alpha} \cdot s(\bar{y})$ ஆகும்.

பெரிய அளவில் அமையும் மாதிரிகட்கு,

$L = 2 \cdot z_{\alpha} \cdot \sigma(\bar{y})$ என்று காணப்படும். குறிப்பிட்ட ஒரு மதிப்பினை $E(L)$-க்குக் பெறும் அளவில், $\left(\text{அல்லது } \frac{E(L)}{\bar{Y}}\text{-க்கு}\right)$, மாதிரி அளவானது நிர்ணயிக்கப்படும். அனுமதிக்கப்படும் பிழை அளவு $\frac{1}{2} \cdot \frac{E(L)}{\bar{Y}}$ என்பதாகும். எனவே, நம்ப எல்லையின் எதிர்பார்க்கும் நீள மதிப்பினைத் தரக்கூடிய அளவில், மதிப்பீட்டின் தொடர்பு படுத்தப்பட்ட தரப்பிழை

$$e = c(\bar{y}) = \frac{d}{z_{\alpha}} \text{ ஆகும்.}$$

(முன்பே பெறப்பட்ட மதிப்பு). இதற்கு, c-ன் மதிப்புத் தெரியப்படல் வேண்டும்.

தீர்மானங்கள்பற்றிய பிரச்னைகளில் மாதிரி அளவு (Sample Size in Decision Problems)

மாதிரி அளவு நிர்ணயித்தலில் ஒரு முக்கிய நிலையினை இப்பொழுது சற்று விரிவாக ஆய்வது தேவையாகிறது. மாதிரித் தேர்வின் மூலம் பெறப்பட்ட முடிவுகளிலிருந்து, நடைமுறையில் ஒரு திட்டவட்டமான தீர்மானம் (Decision) செய்திடும் நிலைகளில், மாதிரி அளவினை முன்கூட்டியே நிர்ணயித்துக் கொள்ளுவது அறிவுடைமையாகும். மாதிரி மதிப்பீடானது, அதிக அளவு பிழையுடன் (error) அமையாது, மிகவும் குறைந்த அளவு பிழையுடன் இருப்பின், மிக நுட்பமானதாகவும், திறனுள்ளதாகவும் காணப்படும்.

z என்ற அளவுப் பிழையானது ஒரு மதிப்பீட்டில் இடம் பெறுவதாகக் கொண்டால், பணத்தின்மூலம் மதிப்பிட்டு, அப்பிழையின் விளைவாக ஒரு முடிவில் ஏற்படும் இழப்பானது (loss) $l(z)$ என்று கணக்கிடப்படும். z-ன் மதிப்பினை முன்கூட்டியே கணக்கிட இயலாவிடினும், அதனைச் சார்ந்த நிகழ் தகவுப் பரவல்

(probability distribution) கண்டறியப்படுவதற்கு, மாதிரிக் கோட்பாடுகள் வழிசெய்கின்றன. இப் பரவலானது, $f(z, n)$ என்று கொள்ளப்படலாம். இது, மாதிரி அளவினை (n) சார்ந்த தாகும். இதன்மூலம், எதிர்பார்க்கும் இழப்பானது.

$$L(n) = \int l(z) \cdot f(z, n)\, dz$$

[z — என்பது வாய்ப்பு மாறி — (Chance variable)] என்பதாக கொடுக்கப்பட்ட மாதிரி அளவிற்கு ஏற்ப வரையறுக்கப்படு கிறது.

இழப்பைக் குறைப்பதே மாதிரித் தேர்வின் அடிப்படை நோக்கமாக அமைவது கவனிக்கத்தக்கது. இனி, அடுத்தது செலவு பற்றிய (cost) நிலையாகும். $C(n)$ என்பது n அளவுள்ள மாதிரியில் ஏற்படும் செலவாக இருப்பின்,

$$C(n) + L(n)$$

என்ற மதிப்பைச் சிறும அளவினதாக்கத்தக்க (minimise) மாதிரி அளவினை, அதாவது n-அளவினைத் தேர்ந்தெடுப்பதே நடைமுறையில் கையாளப்படும் வழி முறையாகிறது. இவ்வாறு தெரிவு செய்யப்படும் n-மதிப்பானது, மாதிரியின் உத்தம (Optimum) அளவினையும், பயன்தரத்தக்கத் திட்ப அளவையும் (precision) நிர்ணயிப்பதாகும்.

மிக எளிய நிலையில், $l(z) = \lambda z^2$ என்றிருப்பின் (λ என்பது மாறிலி), பின்வருமாறு காணலாம் :

எதிர்பார்க்கும் இழப்பு

$$L(n) = \lambda \cdot E(z^2)$$

என்று அமைய, எளிய ராண்டம் மாதிரி முறையினைப் பயன் படுத்தி,

$\hat{\bar{Y}}$ என்பது $\bar{Y}$-ன் எளிய மதிப்பீடாகவும், $z = \hat{\bar{Y}} - \bar{Y}$ எனவும் கொண்டு,

$$L(n) = \lambda \cdot V(\hat{\bar{Y}}) = \frac{\lambda S^2}{n} - \frac{\lambda S^2}{N}$$

$$\left\{ S^2 = \frac{1}{N-1} \sum_{1}^{N} (Y_i - \bar{Y})^2 \right\}$$

என்றவாறு பெறப்படும்.

எளிய செலவுச் சார்பினை

$$C(n) = C_0 + C_1 \quad n$$

என்று எடுத்துக் கொண்டு, (C_0 = மாறாச் செலவு), வகை நுண் கணித முறைப்படி, வகையிட்டு (differentiating), மாதிரி அளவானது.

$$n = \sqrt{\frac{\lambda S^2}{C_1}}$$

என்று தீர்க்கப்படுகிறது. { இது செலவு, மற்றும் இழப்பின் கூட்டு மதிப்பைச் சிறும அளவினதாக்கத்தக்கது. }

குறிப்பு : செலவானது, n-ல் நேர்கோட்டுத் தொடர்புடனும், மதிப்பீட்டு மாறுபாட்டு அளவு n-க்கு எதிர் விகிதத்திலும் இருப்பின், மேற்கண்ட ஆய்வு முடிவு பயன்படும்.

5. படுகை மாதிரி முறை

(Stratified Sampling)

[அறிமுகம் — மதிப்பீடுகள் — எடுத்துக் காட்டுகள் மற்றும் பயிற்சிகள் — மாறுபாட்டு அளவைகள் — எடுத்துக்காட்டுகளும் பயிற்சிகளும் — மாதிரிப் பங்கீடுகள் — (உத்தமப்பங்கீடு ; வீதப் பங்கீடு ; நேமன் பங்கீடு.) — படுகைப்படுத்திய சமவாய்ப்பு முறை — படுகை முறை, சமவாய்ப்புமுறை — ஓர் ஒப்புமை — வீத அளவுபற்றிய படுகைச் சமவாய்ப்பு முறை.]

அறிமுகம் : ஒரு முழுமைத் தொகுதியானது பல பிரிவுகளாக (Strata) அல்லது படுகைகளாகப் பகுக்கப்பட்டு. ஒரு பிரிவினின்றும் அல்லது படுகையினின்றும் ஒன்றுக்கொன்று தொடர்பின்றி (Independent) மாதிரி தேர்ந்தெடுக்கப்பட்டு, இவ்வாறு பெறப்பட்ட மாதிரிகளை ஆய்வுக்குப் பயன்படுத்தும் வழிமுறையானது படுகை மாதிரித் தேர்வுமுறை எனப்படும். கருதப்பட்ட இயல்பின் படுகைவாரியான (Stratum-wise) மதிப்பீடுகள் அனைத்தையும் தக்கமுறையில் இணைத்துப் பெறப்படும் மதிப்பானது முழுமைத் தொகுதி இயல்பின் மதிப்பீடாக (Estimator) கொள்ளப்படும். சில உட்பிரிவுகளைக் கருதும் நிலையில் அதிக திட்பம் தேவைப்படும்போது. முழுமைத் தொகுதியானது இவ்வாறு பிரிக்கப்படுவது பலனளிக்கிறது. ஆய்வுப்பணிகட்கு இம்முறை மிகவும் துணை செய்கிறது. வெவ்வேறு சூழ்நிலைகளில் அமைந்திருக்கும் மக்கட் தொகுதிகள்பற்றிய விவரங்களை இம்முறையின் மூலம் நன்கு மதிப்பிட்டு அறிய இயலும்.

கருதப்படும் படுகைகள் அனைத்தும், ஒவ்வொன்றும் தம்முள் ஒரேபடித்தானவையாக (Homogeneous) அமைதல் இன்றியமையாதது. அந் நிலையில் ஒவ்வொரு படுகையிலும் கிடைக்கும் விவரங்களின் தன்மைக்கேற்ப, மாதிரிமுறை மற்றும் மதிப்பீட்டு முறைகளும் அமைந்து, படுகைகளின் இடைப்பட்ட மாறுபாடு காணப்படும். பொதுவாக, ஒவ்வொரு படுகைக்கும் தக்கவாறு மாதிரி அளவினைப் பங்கீடு செய்யும் பிரச்சினையானது (Allocation of size), மாதிரிமுறை சார்ந்த மாறுபாட்டினையும், செலவையும் குறைக்கத்தக்க விதத்தில் அமையும். அதாவது, பலபடித்தான (Heterogeneous) முழுமைத் தொகுதியை ஒரேபடித்தான

(Homogeneous) உட்பிரிவுகளாகப் பகுத்து, மாறுபாடு பற்றிய ஆய்வு நடத்தப்படும்.

எடுத்துக்காட்டாக, ஒரு நகரில் காணப்படும் சில்லறை விற்பனையாளர் (Retail Stores) தம் சராசரி விற்பனையினை மதிப்பிடும் பொருட்டு, நகரில் உள்ள பல்வேறு விற்பனைக் கூடங்களையும், அவற்றின் மொத்த (Volume of Sales) அடிப் படையில் படுகைகளாகப் பிரித்து, பல்வேறு படுகைகளி லிருந்தும் (Strata) சமவாய்ப்பு அடிப்படையில் மாதிரிகளைத் தெரிவுசெய்து, அவற்றிலிருந்து மதிப்பீடுகள் கணக்கிடப்படும். இத்தகைய கலப்பு வழி முறையானது, படுகை ராண்டம் மாதிரி முறை (Stratified Random Sampling) எனப்படும்.

மக்களின் (அல்லது பொருட்களின்) தன்மைக்கு ஏற்றபடி படுகைகளைப் பிரிக்கலாம். தொழில் அடிப்படையில், வருமான அடிப்படையில், மொழி அடிப்படையில், பிரிவின் அடிப்படையில் என்றவாருகப் படுகைகள் அமைக்கப்படும். ஒவ்வொரு படுகையும் ஒரு தனி முழுமைத் தொகுதியாகக் கருதப்படுவது ஆய்வுக்கு எளிது. மேற்பார்வையிட்டு, பெறும் விவரங்களின் நம்பகத் தன்மையினை உறுதி செய்திடவும் இயலுகின்றது.

எனவே, மதிப்பீட்டுத் திட்பம், மாறுபாட்டு அளவு, ஆய்வுச் செலவு ஆகியவற்றை நன்கு கருத்திற்கொண்டு, படுகை மாதிரிமுறை செயல்படுத்தப்படுகின்றது. ஒவ்வொரு படுகையின் மதிப்பீடும் தனிக் கவனம் செலுத்தப்பட வழி ஏற்படுகிறது.

இந் நியதியின் ஆய்வில், நாம் கருத முற்படுவது படுகை மாதிரி மதிப்பீடுகளின் இயல்புகளைப் பற்றியும், மிகுந்த திட்ப அளவினைப் பெறத் தக்க விதத்தில் படுகை மாதிரி அளவுகளைத் தீர்மானித்தல் பற்றியும் (அதாவது மாதிரிப் பங்கீடு பற்றியும்) அமைகின்றது.

நடைமுறையில் இத்தகைய வழிமுறையானது பெரிதும் பயன் தருவதை ஒரு சில உதாரணங்களில் காணலாம். அரசு மேற்கொள்ளும் பல்வேறு மாதிரி ஆய்வுகளிலும் (உதாரணம் : NSS ஆய்வுகள்) இம் முறையானது பெரிதும் பின்பற்றப்படுவதொன்ருகும்.

(1) ஒரு நிலப்பரப்பில் பயிரிடப்படும் பயிர் இனங்களுக்குத் தக்கபடி நிலக் கூறுகளைப் பல படுகைகளாகப் பிரித்து ஆய்வு மேற்கொள்ளப்படுகிறது. பயிர் விளைச்சல் ஆய்வுகளில் (Crop-cutting Experiments) இத்தகைய வழிமுறை பின்பற்றப்படுகிறது.

(2) ஒரு மாவட்ட நிலப்பரப்பை நகர்ப்புறம் மற்றும் கிராமப் புறம் என்ற அளவில் பிரித்து, ஒவ்வொன்றிலும் துணைத் தொகுதிகள் (blocks) பிரிக்கப்பட்டு, விரிவான அளவில் ஆய்வு மேற்கொள்ளப்படுகிறது.

(3) பல்வேறு தொழில் நிறுவனங்களில் ஆய்வு மேற்கொள்ள வசதி செய்யும் வகையில், அவற்றின் உற்பத்தி அளவிற்கு ஏற்ப (Production) அல்லது சொத்துக்களின் அளவிற்கு ஏற்ப (Capital), அல்லது பணியாளர் எண்ணிக்கைக்கு ஏற்ப அவற்றைப் பல படுகைகளாகப் பிரிப்பது எளிய ஆய்வுக்கு வழி செய்கிறது.

(4) மக்களது வாழும் நிலைபற்றிய பொருளாதார மற்றும் சமூக நிலை குறித்து மேற்கொள்ளப்படும் ஆய்வுகளில் அவர்களை வருமான அடிப்படையில் வெவ்வேறு தொகுதிகளாகப் பிரிப்பது இன்றும் நடைமுறையில் கையாளப்படும் வழியாகும்.

இதேபோல், பல்வேறு நிலைகளிலும், படுகை மாதிரிமுறை ஆய்வு ஒரு தனி இடம் பெற்று விளங்கிடக் காணலாம். இனி, கணக்கீட்டு முறைகள்பற்றி விளக்கமாகப் பகுப்பாய்வு நடத்த முற்படுவோம்.

மதிப்பீடுகள்

முழுமைத்தொகுதி மொத்தம் மற்றும் சராசரி (Population Total & Mean)

N அளவுடைய முழுமைத் தொகுதியிலிருந்து முறையே N_1, $N_2 \ldots\ldots N_L$ என்ற அளவுகள் உடைய L துணைத்தொகுதிகள் அல்லது படுகைகள் (Sub-populations) எடுக்கப்படுகின்றன.

$\therefore N_1 + N_2 + \ldots\ldots + N_L = N$ ஆகும்.

பொதுவாக N_h என்பது h-ஆவது படுகையில் உள்ள கூறுகளின் அளவாகும் (Units).

பின்பு, பல்வேறு படுகைகளினின்றும். ஒன்றுக்கொன்று தொடர்பற்ற விதத்தில் (Independent); முறையே n_1, $n_2 \ldots\ldots n_L$ அளவுகள் கொண்ட மாதிரிகள் தேர்ந்தெடுக்கப்படுகின்றன.

$\therefore n_1 + n_2 + \ldots\ldots + n_L = n$ ஆகிறது.

இதில் பொதுவாக n_h என்பது h ஆவது படுகையினின்று தேர்ந்தெடுக்கப்படும் மாதிரியின் அளவினைக் குறிப்பதாகும்.

Y_{hi} என்பது முழுமைத் தொகுதியில். i-ஆவது உறுப்பின் (Unit) மதிப்பைக் குறிப்பதாகக் கொண்டு, h என்னும் படுகையின் முழுமைத் தொகுதி மொத்தம்

$$Y_h = \sum_{i=1}^{N_h} Y_{hi}$$

என்று பெறப்படும்.

இதனின்று, முழுமைத் தொகுதியில் h-ஆவது படுகை சராசரி

$$\bar{Y}_h = \sum_{1}^{N_h} \frac{Y_{hi}}{N_h} = \frac{Y_h}{N_h}$$

என்றாகும்.

$$\text{மாதிரிச் சராசரி; } \bar{y}_h = \frac{\sum_{1}^{n_h} y_h}{n_h}$$

என்று (h-ஆவது படுகைக்கு) பெறப்படுகிறது.

வெவ்வேறு தனித்தனிப் படுகைகளின் மொத்தங்களை (Y_h, $h = 1, 2 \ldots\ldots L$) கூட்டிப் பெறப்படும் முழுமைத் தொகுதி மொத்தத்தையும், அதனின்று கிடைக்கும் முழுமைத் தொகுதி சராசரியினையும் (Population of Mean) பின்வருமாறு எழுதுகிறோம்:

அதாவது

$$Y = Y_1 + Y_2 + \ldots\ldots + Y_L$$

எனவும்

$$\bar{Y} = \frac{Y}{N} = \frac{Y_1 + Y_2 + \ldots\ldots + Y_L}{N_1 + N_2 + \ldots\ldots + N_L}$$

எனவும், முழுமைத் தொகுதியில் பெறப்படுகின்றன.

இம் மதிப்புகளை மாதிரிகளின் மூலமாக மதிப்பிட்டுக் கணிக்கலாம்.

பல படுகைகளின் மொத்தங்களின் (Totals) மதிப்புக்களைக் கூட்டிப் பெறப்படும் மதிப்பீடானது, முழுமைத் தொகுதி மொத்தத்தின் மதிப்பீடாகும்.

மதிப்பீட்டு அளவைகள் (மாதிரியின் அடிப்படையில்) பற்றிய சில தன்மைகளைக் காண்போம் :

முழுமைத் தொகுதிச் சராசரியினை மதிப்பிடப் பொருத்தமான படுகை மாதிரி மதிப்பீட்டு அளவையாவது

$$\bar{y}_{st} = \frac{\sum_{h=1}^{L} N_h \bar{y}_h}{N}$$

என அமையும். இதில் $N = N_1 + N_2 + \ldots\ldots + N_h$.
(*st* என்ற அடிக்குறி (Suffix), படுகை மாதிரி எடுத்தலைக் குறிக்கும்).

$\bar{y}_{st}$ எனும் மதிப்பு, மாதிரிச் சராசரியின் மதிப்பல்ல. ஏனெனில்

$$\bar{y} = \frac{\sum_{h=1}^{L} n_h \bar{y}_h}{n}$$

என்பதுதான் மாதிரியின் சராசரியாகும்.

அதாவது, $\bar{y}_{st}$-யானது மதிப்பிடப் பெறுகையில். ஒவ்வொரு படுகையின் அல்லது பகுதியின் மதிப்பும் அதற்கேற்ற $\frac{N_h}{N}$ என்னும் காரணியால் அல்லது எடை அளவால் (weight) பெருக்கப்படும்.

$\frac{N_h}{N} = \frac{n_h}{n}$ அல்லது $\frac{n_h}{N_h} = \frac{n}{N}$ என்று இருப்பின் $\bar{y}_{st} = \bar{y}$ என்று காணலாம்.

அதாவது மாதிரிமுறைப் பின்ன (Sampling fraction) அளவானது ஒவ்வொரு படுகையிலும் மாறாத அளவாக (மாறிலி) இருத்தல் அவசியம். இது, n_h-ன் 'வீதப் பங்கீட்டுமுறை' (Proportional Allocation) என்று கருதப்படும் நிலை ஆகும்.

தேற்றம் 1.

ஒவ்வொரு படுகையிலும் (stratum), பெறப்படும் $\bar{y}_h$ என்ற மாதிரிச் சராசரியானது, பிறழ்ச்சியற்ற மதிப்பீடாக அமையுமானால், $\bar{y}_{st}$ என்ற மதிப்பானது, $\bar{Y}$ என்ற முழுமைத் தொகுதிச் சராசரியின் பிறழ்ச்சியற்ற (unabiased) மதிப்பீடாகும்.

நிரூபணம் : $E(\bar{y}_{st}) = \bar{Y}$ என்று நிறுவுதல் வேண்டும்.

வரையறைப்படி, $\bar{y}_{st} = \sum_{h=1}^{L} \frac{N_h \cdot \bar{y}_h}{N}$ ஆகும்.

எனவே. $E(\bar{y}_{st}) = E\left[\sum_{h=1}^{L} \frac{N_h \cdot \bar{y}_h}{N}\right]$

$$= \sum_{h=1}^{L} \frac{N_h \cdot E(\bar{y}_h)}{N}$$

$$= \sum_{h=1}^{L} \frac{N_h \cdot \bar{Y}_h}{N} \quad \left[\because \; E(\bar{y}_h) = \bar{Y}_h\right] \quad \ldots\ldots \text{(I)}$$

படுகை மாதிரி மதிப்பீடுகள் பிறழ்ச்சியற்றிருப்பின், இவ்வாறு பெறுகிறோம்.

ஆனால், முழுமைத் தொகுதி சராசரியான $\bar{Y}$-ஐப் பின் வருமாறு எழுதலாம்:

$$\bar{Y} = \frac{\sum_{h=1}^{L} \sum_{i=1}^{N_h} y_{hi}}{N}$$

(h என்பது கருதப்படும் படுகை (stratum); i என்பது உறுப்பு எண்ணிக்கை).

$$\text{அதாவது,} \quad \bar{Y} = \frac{\sum_{h=1}^{L} N_h \cdot \bar{Y}_h}{N} \quad \text{(II)}$$

$$\left(\because \; \bar{Y}_h = \frac{\sum_{i} y_{hi}}{N_h}\right)$$

(I), (II) இரண்டையும் இணைத்து,

$$E(\bar{y}_{st}) = \bar{Y}$$

என நிறுவலாம்.

துணைமுடிவு :

$\bar{y}_h$ என்ற மதிப்பீடானது, $\bar{Y}$-ன் பிறழ்ச்சியற்றப் மதிப்பீடாக (uubiased estimate) ஒவ்வொரு படுகையின் சமவாய்ப்பு மாதிரியிலும் அமைகையில், $\bar{y}_{st}$ என்ற இணைப்பு மதிப்பீடானது, $\bar{Y}$-ன் பிறழ்ச்சியற்ற மதிப்பீடாக, படுகைமுறை மாதிரி எடுத்தலில் பெறப்படும்.

எடுத்துக்காட்டு 1 : 6 மாணவர்கள் உள்ள முழுமைத் தொகுதியின், இரண்டு படுகைகளில் காணப்பட்ட y-மதிப்புக்கள் வருமாறு :—

படுகை 1 : 2, 4, 6 $(N_1 = 3)$

படுகை 2 : 8, 12, 16 $(N_2 = 3)$

$N = N_1 + N_2 = 3+3 = 6.$

இரு படுகைளின் கூட்டுமொத்தங்கள் (Strata Totals)

$Y_1 = 2 + 4 + 6 = 12$;

$Y_2 = 8 + 12 + 16 = 36$

இவற்றின் சராசரிகள் :—

$$\bar{Y}_1 : \frac{y_1}{N_1} = \frac{12}{3} = 4$$

$$\bar{Y}_2 : \frac{y_2}{N_2} = \frac{36}{3} = 12$$

$$\therefore \bar{Y} = \frac{Y_1 + Y_2}{N_1 + N_2} = \frac{12 + 36}{3+3} = \frac{48}{6} = 8$$

ஒவ்வொரு படுகையிலிருந்தும் எடுக்கப்படும் எல்லா மாதிரிகளையும் கருதலாம். ஒவ்வொன்றிலிருந்தும், $\binom{3}{2} = 3$ மாதிரிகளை (திரும்ப வைக்காத முறையில்) எடுக்கலாம். ∴ மொத்தமாக, 9 மாதிரிகள் கிடைக்கின்றன.

படுகை 1	படுகை 2	y_1	y_2	$\bar{y}_1$	$\bar{y}_2$	$N_1\bar{y}_1$	$N_2\bar{y}_2$
2, 4	8, 12	6	50	3	10	9	30
	8, 16		24		12	9	36
	12, 16		28		14	9	42
2, 6	8, 12	8	20	4	10	12	30
	8, 16		24		12	12	36
	12, 16		28		14	12	42
4, 6	8, 12	10	20	5	10	15	30
	8, 16		24		12	15	36
	12, 16		28		14	15	42

முதல் மாதிரியில், (2, 4, 8, 12) என்ற மதிப்புகளில்; (2, 4) என்பவை முதல்படுகையினின்றும், (8, 12) என்பவை இரண்டாம் படுகையினின்றும் வருபவை.

∴ $\bar{y}_1 = 2 + 4 = 6 ; \quad \bar{y}_2 = 8 + 12 = 20,$

∴ $\bar{y}_1 = 3 ; \quad \bar{y}_2 = 10.$

∴ முதல் மாதிரியின் மூலம் பெறப்படும் மொத்த மதிப்பு மதிப்பீடானது $\hat{Y}_{st} = N_1 \bar{x}_1 + N_2 \bar{x}_2 = 9 + 30 = 39.$ இதன் மூலம் பெறப்படும் சராசரியின் மதிப்பீடு (Estimate of the Mean),

$$\bar{y}_{st} = \frac{\hat{Y}_{st}}{N} = \frac{39}{6} = 6{\cdot}5 \text{ ஆகிறது.}$$

அதாவது $N_1\bar{y}_1$, $\bar{y}N_2\bar{y}_2$ மதிப்புகளைக் கண்டுபிடித்து, ஒவ்வொரு மாதிரிக்கும், அவற்றின் கூட்டுத் தொகையைக் கணக்கிட்டு, அவற்றின்மூலம் படுகை மாதிரியின் சராசரி $\bar{y}_{st}$ யானது மதிப்பிடப்பெறுகிறது.

மேலும்,

$$E(\bar{y}_{st}) = \frac{N_1 \bar{y}_1 + N_2 \bar{y}_2}{N_1 + N_2} \text{ (எல்லா மாதிரிகளையும் கருதி)}$$

$$= \frac{1}{9}\left[\frac{39}{6} + \frac{45}{6} + \ldots\ldots + \frac{57}{6}\right]$$

$$\therefore \quad E(\bar{y}_{st}) = 8 = \bar{Y}_{st}.$$

எனவே $\bar{y}_{st}$ ஆனது $\bar{Y}_{st}$-யின் பிறழ்ச்சியற்ற மதிப்பீடாக அமைகிறது.

பயிற்சி

$\hat{Y}_{st}$ யானது Y-யின் பிறழ்ச்சியற்ற மதிப்பீடென நிறுவுக.

பயிற்சி

(1) இரு படுகைகளின் முழுமைத் தொகுதிக் கூறுகளாவன:

முதல் படுகை : 20, 40, 60

இரண்டாம் ,, : 9, 12, 18, 21.

(*i*) Y_1, Y_2, $\bar{Y}_1$, $\bar{Y}_2$ மேலும் $\bar{Y}$ ஆகியவற்றை கணக்கிடுக.

(*ii*) $n_1 = 2$, $n_2 = 3$ என்றும் அளவுகளுடைய எல்லா வழி மாதிரிகளையும் எடுத்து மேற்கண்டவற்றை மதிப்பீடு செய்க.

(*iii*) எடுத்த மாதிரிகளின் அடிப்படையில்,

$E(\bar{y}_{st}) = \bar{Y}_{st}$ எனவும்

$E(\bar{y}_{st}) = \bar{Y}$ எனவும் நிறுவுக.

(2) சில விற்பனைக் கடைகளின் (shops) (இடப்) நிலப் பரப்பு (floor space) கணிக்கப்படுவதற்காக எடுக்கப்பட்ட மாதிரி மதிப்புகள் கொடுக்கப்படுகின்றன:

படுகை (அல்லது பிரிவு)	N_h	n_h	x_h
சிறியது	110	20	4000 சதுர அடி
நடு அளவினது	60	15	10500 ,,
பெரியது	30	10	60000 ,,

(*i*) $\bar{Y}_1$, Y_2 & $\bar{Y}$ ஆகியவற்றைக் கணக்கிட்டு, அவற்றின் மாதிரி மதிப்பீடுகளைப் பெறுக.

(*ii*) $\bar{Y}$ அளவை மதிப்பிட்டு $E(\bar{y}_{st}) = \bar{Y}$ என்று நிறுவுக.

(3) படுகை சமவாய்ப்பு மாதிரித் தேர்வில் பின்வரும் பட்டியல் கொடுக்கப்படுகிறது. முழுமைத் தொகுதி சராசரிக்கும், அச் சராசரியின் மாதிரி முறை சார்ந்த மாறுபாட்டுக்கும் மதிப்பீடுகள் காண்க.

படுகை	N_h	S_h^2	n_h	$\bar{y}_h$
1	72	44	9	14·2
2	65	27	8	11·6
3	53	35	7	16·1

{குறிப்பு : S_h^2 கொடுக்கப்பட்டதால்,

$$V(\bar{y}_{st}) = \frac{1}{N^2} \cdot \Sigma N_h^2 \cdot \frac{N_h - n_h}{N_h} \cdot \frac{S_h^2}{n_h}$$

என்றபடி மதிப்பிடப் பெறலாம்.}

(4) ஒரு பயிர் விளைச்சல் நிலப் பரப்பு பற்றிய ஆய்வினில், எடுத்துக் கொள்ளப்பட்ட மாவட்டத்தில் உள்ள 186 கிராமங்கள் 4 படுகைகளாகப் பிரிக்கப்பட்டு, ஒவ்வொரு படுகையினின்றும் சமவாய்ப்பு மாதிரிகள் கருதப்பட்டு, வீதப் பங்கீட்டு அடிப் படையில் எடுக்கப்படும் கிராமங்களில் காணப்பட்ட பரப்புகள் விவரம் கொடுக்கப்பட்டுள்ளது.

படுகை	N_h	n_h	மாதிரிக்கிராமங்களில் விளை நிலப் பரப்பு
1	72	8	14, 12, 8, 11, 12, 10, 13, 16
2	53	5	27, 20 21, 22, 30
3	35	4	36, 47, 52, 61
4	26	3	92, 105, 82

பயிர் விளையும் மொத்த நிலப் பரப்பினை மதிப்பிட்டு, அதன் தரப் பிழை அளவினையும் நிர்ணயிக்க.

(B.Sc April, 1973)

படுகை மாதிரியில் மாறுபாட்டு அளவைகள்

$\bar{y}_{st}$-யின் திட்பத்தினை முடிவு செய்தற்பொருட்டு, அதனு டைய மாதிரி முறைசார் மாறுபாட்டினை (Sampling variance), அதாவது $V(\bar{y}_{st})$ மதிப்பைகாணல் வேண்டும். இதில் $V(\bar{y}_{st})$ ஆனது, இருவகையாக பிரிக்கப்படும். முதல் பகுதி, கருதப் பட்ட ஒவ்வொரு படுகையிலும் இருக்கும் உள்ளீட்டு மாறுபாடு

(within stratum); மற்றொன்று பல படுகைகளுக்கிடையே உள்ள இடையீட்டு மாறுபாடு (between strata). இவ்விரண்டினை நன்கு அறிந்துகொள்வது, படுகைப் படுத்தும் மாதிரி முறையினை நன்கு பின்பற்றவும், திட்டமிடல் செவ்வனே அமையவும் வழி செய்கிறது.

பல்வேறு மாறுபாட்டு அளவைகளை முதலில் வரையறை செய்கிறோம்:

படுகை மாறுபாடு (ஒரு படுகையில் அல்லது பிரிவில் கருதப்படும் மாறுபாட்டு அளவு).

$$\sigma_h^2 = \frac{1}{N_h} \sum_{1}^{N_h} (Y_{hi} - \bar{Y}_h)^2$$

மற்றொரு வரையறை :

$$S_h^2 = \frac{1}{N_h - 1} \sum_{1}^{N_h} (Y_{hi} - \bar{Y}_h)^2$$

S_h^2 என்பது உண்மை மாறுபாட்டு அளவாகக் கருதப்படும் (true variance).

ஆனால் முழுமைத் தொகுதியை அப்படியே கருதிப் பெறப்படும் மாறுபாட்டு அளவு

$$\sigma^2 = \frac{1}{N} \sum_{1}^{L} \sum_{i=1}^{N_h} (Y_{hi} - \bar{Y})^2$$

என்பது வேறாகும்.

மேலும்,

$$\sigma^2 = \frac{1}{N} \sum_{h}^{L} \sum_{i}^{N_h} (Y_{hi} - \bar{Y})^2$$

$$= \frac{1}{N} \sum_{h} \sum_{i} [(Y_{hi} - \bar{Y}_h) + (\bar{Y}_h - \bar{Y})]^2$$

$$= \frac{1}{N}\left[\sum_{h}\sum_{i}(Y_{h} - Y)^2 + \sum_{h=1}^{L} N_h(\bar{Y}_h - \bar{Y})^2\right]$$

$$= \frac{1}{N}\sum_{h} N_h \cdot \sigma_h^2 + \frac{1}{N}\sum_{h} N_h(\bar{Y}_h - \bar{Y})^2$$

∴ மொத்த மாறுபாடானது, படுகைகளின் உள்ளீட்டு மாறுபாட்டினையும், இடையீட்டு மாறுபாட்டையும் கூட்டிப் பெறுவதாகும்.

தேற்றம் 2 :

படுகை மாதிரிமுறையில் (Stratified Sampling) $\bar{y}_{st}$ என்ற சராசரி மதிப்பீட்டின் மாதிரிமுறைசார் (Sampling) மாறுபாடானது

$$V(\bar{y}_{st}) = \frac{\sum_{h=1}^{L} N_h^2 . V(\bar{y}_h)}{N^2}$$

ஆகும். இதில் $V(\bar{y}_h) = E(\bar{y}_h - \bar{Y}_h)^2$ ஆக அமையும். இதற்கு இரு நிபந்தனைகளாவன:

ஒன்று: $\bar{y}_h$ என்ற மாதிரி அளவையானது. $\bar{Y}_h$ என்ற முழுமைத் தொகுதி அளவையின் பிறழ்ச்சியற்ற மதிப்பீடாதல் வேண்டும்.

இரண்டு: பல்வேறு பிரிவுகளிலிருந்து அல்லது படுகைகளிலிருந்து மாதிரிகளை ஒன்றோடொன்று தொடர்பற்ற நிலையில் எடுக்கவேண்டும்.

நிருபணம் :

$\bar{y}_{st}$ மற்றும் $\bar{Y}$ இவற்றிற்கு இடைப்பட்ட வேறுபாட்டினை

$$(\bar{y}_{st} - \bar{Y}) = \frac{\sum_{h=1}^{L} N_h . \bar{y}_h}{N} - \frac{\sum_{h=1}^{L} N_h . \bar{Y}_h}{N}$$

$$= \frac{\sum_{h=1}^{L} N_h (\bar{y}_h - \bar{Y}_h)}{N}$$

என்று எழுதலாம்.

$(\bar{y}_{st} - \bar{Y})$ என்ற மதிப்பீட்டுப் பிழையானது (error in the estimate), வெவ்வேறு படுகைகளில் அல்லது பகுதிகளில் உள்ள மதிப்பீட்டுப் பிழைகளின் எடையிட்ட அல்லது நிறையிட்ட (weighted) சராசரியாகக் காணலாம்.
எனவே,

$$(\bar{y}_{st} - \bar{Y})^2 = \frac{\sum_{h=1}^{L} (\bar{y}_h - \bar{Y})^2}{N^2}$$

$$+ \; 2 \; \frac{\sum_{h \neq j} N_h \; N_j \; (\bar{y}_h - \bar{Y}_h) \; (\bar{y}_j - \bar{Y}_j)}{N^2}$$

வலப்புறம் இரண்டாவது உறுப்பு, (h, j) என்ற எல்லா இணைகளுக்கும் (pairs) அமைகிறது.

மேற்கண்ட மதிப்பினை, எல்லா மாதிரிக் கூறுகளையும் கருதிச் சராசரியிடுகிறோம்.

குறுக்காகப் பெருக்கிய தொடர்களில் (cross products) h படுகையில் கருதிய மாதிரியினை நிலையாகக் (Fixed) கருதி, j என்ற படுகையில் எடுக்கப்படும் எல்லா மாதிரிகளையும் கருதிய சராசரியினைக் காணும்போது, அத்தகைய குறுக்குப் பெருக்கல் உறுப்பானது மதிப்பில் பூச்சியமாகிறது. (ஏனெனில், $\bar{y}_j$ என்பது பிறழ்ச்சியற்றதாகக் கருதப்படுவதால், $(\bar{y}_j - \bar{Y}_j)$-ன் சராசரி மதிப்பு பூச்சியமாகும்.)

எனவே,

$$V(\bar{y}_{st}) = E[\bar{y}_{st} - \bar{Y}]^2$$

$$= \sum_{h=1}^{L} \frac{N_h^2 \; E(\bar{y}_h - \bar{Y}_h)^2}{N^2}$$

$$\therefore V(\bar{y}_{st}) = \sum_{h=1}^{L} \frac{N_h^2 \; V(\bar{y}_h)}{N^2}$$ எனக் காணலாம்.

$V(\bar{y}_{st})$ யானது படுகைச் சராசரிகளின் அதாவது $\bar{Y}_h$-ன் மதிப்பீடுகளின் $(\bar{y}_h)$ மாதிரிமுறைகள் மாறுபாட்டு அளவுகளைப் பொறுத்தே நிர்ணயிக்கப்படுகிறது.

பின்குறிப்பு : மேற்கண்ட தேற்ற உண்மையானது, பொதுவாக எந்த முறையிலும் அமையும் படுகை மாதிரி முறைக்குப் பொருந்தியதாகும்.

குறிப்பாக, படுகைகளினின்று எடுக்கப்படும் ராண்டம் மாதிரி முறையில், இத் தேற்றமானது பின்வருமாறு வடிவமைவு பெறும்.

தேற்றம் 3

படுகையாக்கிய சமவாய்ப்பு மாதிரிமுறையில், $\bar{y}_{st}$-யின் மாதிரி சார்ந்த மாறுபாடானது (Sampling Variance)

$$V(\bar{y}_{st}) = \frac{1}{N^2} \sum_{h=1}^{L} N_h (N_h - n_h) \frac{S_h^2}{n_h}$$

என்பதாகும்.

நிரூபணம்:

$\bar{y}_h$ ஆனது $\bar{Y}_h$—ன் பிறழ்ச்சியற்ற மதிப்பீடாக அமைவதால், சென்ற தேற்ற வரையறை பயன்படுத்தப்படும்.

மேலும், எளிய சமவாய்ப்பு மாதிரிமுறையில் நிறுவிய தேற்ற அடிப்படையில், ஒரு குறிப்பிட்ட படுகையினைக் கருதி,

$$V(\bar{y}_h) = \frac{S_h^2}{n_h} \left(\frac{N_h - n_h}{N_h} \right)$$

என்று பெறலாம்.

$\therefore$ முன் நிறுவப்பட்ட உண்மைப்படி,

$$V(\bar{y}_{st}) = \frac{1}{N^2} \sum_{h=1}^{1} N_h^2 . V(\bar{y}_h)$$

என்பதில், $V(\bar{y}_h)$ க்குப் பிரதியிட,

$$V(\bar{y}_{st}) = \frac{1}{N^2} \sum_{h=1}^{L} N_h (N_h - n_h). \frac{S_h^2}{n_h}$$

என்று பெறலாம்.

துணை முடிவுகள் :

(1) மாதிரிமுறை பின்னம் (sampling fraction) பொருட் படுத்தத் தக்க அளவில் அமைந்திடாத நிலையில், (அதாவது $\frac{n_h}{N_h}$ மதிப்புகள் தள்ளத்தக்க அளவில்) சிறியதாக இருப்பின், (ஒவ்வொரு படுகையிலும்),

$$V(\bar{y}_{st}) = \frac{1}{N^2} \sum_{h=1}^{L} \frac{N_h^2 . S_h^2}{n_h}$$

$$= \sum_{h=1}^{L} \frac{W_h^2 S_h}{n_h}$$

ஆக அமைந்திருக்கும். $W_h = \frac{N_h}{N}$ என்ற அளவானது, ஒவ்வொரு படுகையிலும் உள்ள எடை (நிறை) ஆகும். இங்கு *fpc* ஆனது கருதப்படுவதில்லை. *fpc* என்ற எந்தக் காரணியுமே அமைவதில்லை.

(2) **வீத அளவு ஒதுக்கீட்டு முறையில் (Proportional Allocation)**

$n_h = \frac{n}{N} . N_h$ என்று பிரதியிட,

$$V(\bar{y}_{st}) = \frac{1}{N^2} \Sigma N_h (N_h - n_h). \frac{S_h^2}{n_h}$$

என்ற மாறுபாட்டு அளவானது,

$V(\bar{y}_{st}) = \frac{N-n}{Nn} . \Sigma W_h . S_h^2$ என ஆகும். இதில், ஒரே ஒரு *fpc* ஆனது $\left(\frac{N-n}{N}\right)$ ஆகும்.

(3) வீத அளவீட்டு முறையில் மாதிரி கருதப்படுகையில், மேலும் எல்லாப் படுகைகளிலும் (strata) S_w^2 என்ற மாறுபாடு அமையப்பெறின்,

$$V(\bar{y}_{st}) = \frac{S_w^2}{n} \cdot \left(\frac{N-n}{N}\right)$$

என்று எளிதாகப் பெறப்படும்.

தேற்றம் 4

$\hat{Y}_{st} = N\bar{y}_{st}$ என்பது முழுமைத்தொகுதி அளவான (total) Y-ன் மதிப்பீடானல்,

$$V(\hat{Y}_{st}) = V(N.\bar{y}_{st})$$
$$= N^2 V(\bar{y}_{st})$$
$$= \sum_{h=1}^{L} N_h (N_h - n_h) \frac{S_h^2}{n_h}$$

என்ற முடிவு, சென்ற தேற்றத்தின் அடிப்படையில் பெறப்படுகிறது.

மாறுபாட்டு மதிப்பீடுகள் (மாதிரிக்கு) (Sample Estimates of Variances)

ஒவ்வொரு படுகையிலிருந்தும் (Stratum) ஓர் எளிய சமவாய்ப்பு மாதிரி தெரிந்தெடுக்கப்படின், S_h^2 என்ற மாறுபாட்டிற்கு ஒரு பிறழ்ச்சியற்ற மதிப்பீடானது,

$$s_h^2 = \frac{1}{n_h - 1} \sum_{i=1}^{n_h} (y_{hi} - \bar{y}_h)^2$$

என்று அமையும் நிலையில், பின்வரும் தேற்றத்தினை நிறுவுகிறோம்.

தேற்றம்

படுகைப்படுத்திய, சமவாய்ப்பின் அடிப்படையில் அமைந்த மாதிரிமுறையில், $\bar{y}_{st}$ என்ற மதிப்பீட்டின் மாதிரிசார் மாறுபாட்டின் பிறழ்ச்சியற்ற மாதிரி மதிப்பீட்டு அளவையானது (unbiased estimate)

$$v(\bar{y}_{st}) = s^2(\bar{y}_{st})$$

$$= \frac{1}{N^2}\sum_{h=1}^{L} N_h(N_h - n_h)\frac{s_h^2}{n_h}$$

ஆகும். அதாவது, S_h^2 என்பதன் பிறழ்ச்சியற்ற மதிப்பீடான s_h^2 என்பது $V(\bar{y}_{st})$-ல் பிரதியிடப்பட்டது.

கணக்கீட்டு எளிமை கருதி,

$$s^2(\bar{y}_{st}) = \sum_{h=1}^{L}\frac{W_h^2 . s_h^2}{n_h} - \sum_{h=1}^{L}\frac{w_h . s_h^2}{N}$$

என்ற வரையறை பயன்படுத்தப்படும். இதில் வலப்புற இரண்டாம் உறுப்பானது (term) pcf-ன் மூலம் ஏற்படும் மாறுபாட்டுக் குறைவை (reduction) காட்டுகிறது. மேலும், இந்த மதிப்பீட்டினைக் காணுதற்கு, ஒவ்வொரு படுகையிலிருந்தும் குறைந்தது இரண்டு கூறுகளையேனும் (units) தெரிதல் அவசியம்.

நம்பக எல்லைகள் (Confidence Intervals)

முழுமைத் தொகுதிச் சராசரிக்கு : $\bar{y}_{st} \pm z.\ s(\bar{y}_{st})$

,, ,, மொத்தத்திற்கு : $\bar{y}_{st} \pm z.\ N.s(\bar{y}_{st})$

z என்ற காரணி கொடுக்கப்படும் நம்பகத்தன்மை அளவின் அடிப்படையில், இயல்நிலைப் பரவலைக் கருதிப் பெறப்படும். $s(\bar{y}_{st})$ என்ற மதிப்பு, $\bar{y}_{st}$-ன் தரப்பிழை (standard error) அளவாகும்.

இவ் வெல்லைகள்மூலம் கொடுக்கப்படும் நம்பகத்தன்மையில் [பொதுவாக $(1-\alpha)$ % என்று தரப்படும்], முழுமைத் தொகுதி அமையத்தக்க இடைவெளியினை, மாதிரியினின்று பெறுகிறோம். தரப்பிழையானது மதிப்பீட்டின், மாதிரிமுறை சார்ந்த மாறுபாட்டு மதிப்பின் (sampling variance) வர்க்கமூலமாகும்.

எடுத்துக்காட்டு 1

மாறுபாட்டு அளவுகள் மற்றும் அவற்றின் மதிப்பீடுகள் காணுதல்.

$N = 6.$

முதல் படுகை மதிப்புகள் :

(y_1) : 2, 4, 6 $(N_1 = 3)$

இரண்டாம் படுகை மதிப்புகள் :

(y_2) : 8, 12, 16. $(N_2 = 3)$.

y_{1i}	y_{2i}	y_{1i}^2	y_{2i}^2
2	8	4	64
4	12	16	144
6	16	36	256
12	36	56	464

(*i*) முழுமைத் தொகுதி மாறுபாடு:

ஒவ்வொரு படுகைக்கும் (stratum) காணப்படும் தனித்தனி மாறுபாடுகள்:

$$\sigma_h^2 = \frac{1}{N_h} \sum_{i=1}^{N_h} (y_{hi} - \bar{Y}_h)^2$$ என்பது σ_1^2, σ_2^2 இவற்றைக் காண உதவும்.

$$\sigma_1^2 = \frac{1}{N_1} \left[\Sigma y_{1i}^2 - \frac{1}{N_1} (\Sigma y_{1i})^2 \right]$$

$$= \frac{1}{3} \left[56 - \frac{(12)^2}{3} \right] = \frac{8}{3} = 2{\cdot}67$$

$$\sigma_2^2 = \frac{1}{3} \left[464 - \frac{(36)^2}{3} \right]$$

$$= \frac{32}{3} = 10.67$$

மற்றொரு மாறுபாடு :

$$S_h^2 = \frac{1}{N_h - 1} \sum_{1}^{N_h} (y_{hi} - \bar{Y}_h)^2.$$

$$\therefore \quad S_1^2 = \frac{8}{2} = 4 \ \&$$

$$S^2 = \frac{32}{2} = 16.$$

மொத்தத் தெ[redacted]தியின் மாறுபாடு:

$$\sigma^2 = \frac{1}{N} \Sigma \Sigma (y_{hi} - \bar{Y})^2$$

$$= \frac{1}{N} \left[\sum_{1}^{N_1} (y_{1i} - \bar{Y})^2 + \sum_{1}^{N_2} (y_{2i} - \bar{Y})^2 \right]$$

$\left[\bar{Y} \text{ மதிப்பானது } \frac{Y_1 + Y_2}{N_1 + N_2} = 8{\cdot}0 \right]$

$$\sum_{1}^{N_1} (y_{1i} - \bar{Y})^2 = \sum_{N_1} y_{1i}^2 - 2\bar{Y} \sum_{N_1} y_{1i} + N_1 \bar{Y}^2$$

$$= 56 - (2 \times 8 \times 12) + 3 \cdot (8)^2$$

$$= 56$$

$$\therefore \sum_{1}^{N_2} (y_{2i} - \bar{Y})^2 = \sum_{N_2} y_{2i}^2 - 2\bar{Y} \sum_{N_2} y_{2i} + N_2 \bar{Y}^2$$

$= 80$ என்று பெறலாம்.

இவற்றிலிருந்து,

$$\sigma^2 = \frac{1}{6}(56 + 80)$$

$$= \frac{136}{6} = 22{\cdot}67$$

என்று பெறப்படும்.

வரையறைப்படி,

$$\sigma^2 = \frac{1}{N} \sum_{1}^{N} (y_i - \bar{Y})^2$$

$$= \frac{1}{6}\,[\,(2-8)^2 + (4-8)^2 + \ldots\ldots + (16-8)^2\,]$$

$$= \frac{136}{6} = 22{\cdot}67$$

என்று கணக்கிடலாம். இரண்டும் ஒரே மதிப்பைத் தருகின்றன.

எடுத்துக்காட்டு 2

1, 2, 3, 4, 5, 6, 7 எனும் எண்களை, (1, 2, 3, 4) என்ற எண்கள் முதல் படுகையிலும், (5, 6, 7) ஆகியவை இரண்டாம் படுகையிலும் அமையும்படியாகப் பிரிவாக்குகிறோம்.

முதல் படுகை	இரண்டாம் படுகை
$N_1 = 4$	$N_2 = 3$
$Y_1 = 10\ ;\ \ \bar{Y}_1 = 2{\cdot}5$	$Y_2 = 18\ ,\ \ \bar{Y}_2 = 6$
$(y_{1i} - \bar{Y}_1)^2 :\ 2{\cdot}25,\ 0{\cdot}25,\ 0{\cdot}25,\ 2{\cdot}25$	$(y_{2i} - \bar{Y}_2)^2 :\ 1,\ 0,\ 1.$
$\sigma_1^2 = \frac{1}{N_1}\Sigma(y_{1i} - \bar{Y}_1)^2$	$\sigma_2^2 = \frac{1}{N_2}\Sigma(y_{2i} - \bar{Y}_2)^2$
$= 1{\cdot}00$	$= \frac{2}{3} = 0{\cdot}67$
$S_1^2 = \frac{1}{N_1-1}\Sigma(\bar{Y}_{1i} - \bar{y}_1)^2$	$S_2^2 = \frac{1}{N_2-1}\Sigma(y_{2i} - \bar{Y}^2)^2$
$= \frac{4}{3} = 1{\cdot}67$	$= \frac{2}{2} = 1.$

$$\sigma^2 = \frac{1}{N}[\Sigma\Sigma(y_{hi} - \bar{Y}_h)^2] = \frac{1}{7}[4+2] = \frac{6}{7} = 0{\cdot}8$$

எடுத்துக்காட்டு 3

$\bar{y}_{st}$-யின் மாதிரிசார் மாறுபாடு (sampling variance) மற்றும் அதன் மதிப்பீடு காணுதல்.

இரு படுகைகள் அமைந்த முழுமைத் தொகுதியின் மதிப்புகள்:

படுகை I : 2, 4, 6

படுகை II : 8, 12, 6.

$n = 2$ என்ற அளவுகளில், எல்லா மாதிரிகளையும் ஒவ்வொரு படுகையிலிருந்தும் தெரிவுசெய்து, அவற்றின் அடிப் படையில் $V(\bar{y}_{st})$ மற்றும் அதன் மாதிரி மதிப்பீட்டைக் காண்க.

மாதிரிகள்பற்றிய பட்டியல்

படுகை 1	படுகை 2	$\hat{Y}_{st}$	$\bar{y}_{st}$	$(\bar{y}_{st} - \bar{Y})$	$(\bar{y}_{st} - \bar{Y})^2$
2, 4	8, 12	39	$\frac{39}{6}$	$-\frac{9}{6}$	$\frac{81}{36}$
	8, 16	45	$\frac{45}{6}$	$-\frac{3}{6}$	$\frac{9}{36}$
	12, 16	51	$\frac{51}{6}$	$\frac{3}{6}$	$\frac{9}{36}$
2, 6	8, 12	42	$\frac{42}{6}$	$-\frac{6}{6}$	1
	8, 16	48	$\frac{48}{6}$	0	0
	12, 16	54	$\frac{54}{6}$	$\frac{6}{6}$	1
4, 6	8, 12	45	$\frac{45}{6}$	$-\frac{3}{6}$	$\frac{9}{36}$
	8, 16	51	$\frac{51}{6}$	$\frac{3}{6}$	$\frac{9}{36}$
	12, 16	57	$\frac{57}{6}$	$\frac{9}{6}$	$\frac{81}{36}$

$\hat{Y}_{st}$ என்பது ஒவ்வொரு Y மதிப்புக்கும் (மொத்தம்) உள்ள மாதிரி மதிப்பீட்டு அளவை ஆகும்.

உதாரணமாக, முதல் மாதிரி (2, 4, 8, 12)-ல்,

$$\bar{y}_1 = 3 \qquad \therefore \quad N_1 \bar{y}_1 = 9$$
$$\bar{y}_2 = 10; \qquad N_2 \bar{y}_2 = 30$$

எனவே, $\hat{Y} = 30 + 9 = 39$ ஆகிறது. இதேபோல், $\hat{Y}_{st}$-யின் மற்ற மதிப்புகள் பெறப்படும்.

$\bar{Y}$-ன் மதிப்பு 8 என்று பெறப்பட்டது. (எ·டு. : 1).

$$\therefore \quad V(\bar{y}_{st}) = \frac{1}{9}\left[\frac{81}{36} + \ldots\ldots + \frac{81}{36}\right]$$

$$= \frac{1}{9} \times \left(\frac{270}{36}\right) = \frac{30}{36} = \frac{5}{6} \text{ ஆகும்.}$$

(கடைசிப் பத்தியிலிருந்து).

மேலும்,

$$V(\bar{y}_{st}) = \frac{1}{N^2} \sum_h \frac{(N_h S_h)^2}{n_h} \cdot \frac{N_h - n_h}{N_h}$$

என்ற வரையறைப்படி,

$$V(\bar{y}_{st}) = \frac{1}{N^2}\left[\frac{(N_1 - n_1)}{N_1} \cdot \frac{(N_1 S_1)^2}{n_1} + \frac{(N_2 - n_2)}{N_2} \cdot \frac{(N_2 S_2)^2}{n_2}\right]$$

$$= \frac{1}{36}\left[\frac{(3-2)}{3} \cdot \frac{(9 \times 4)}{2} + \frac{(3-2)}{3} \cdot \frac{(9 \times 16)}{2}\right]$$

$$\therefore \quad \sigma_{\bar{y}_{st}}^{\ 2} = \frac{30}{36} = 0{\cdot}83.$$

∴ இருவழிகளில் கணக்கிடப்பட்ட $\bar{y}_{st}$-ன் மாறுபாட்டு அளவு ஒன்றாக இருப்பதைக் காணலாம்.

$V(\bar{y}_{st})$ - ன் மதிப்பீட்டு அளவை (estimator) காண, S_h^2 மதிப்பிற்கு, அதன் பிறழ்ச்சியற்ற மதிப்பீட்டு அளவினைப் பிரதியிடல் வேண்டும். அதாவது, S_h^2 என்ற மாதிரி மதிப்பீடுகளைக் கணக்கிடவேண்டும்.

மேற்கண்ட எடுத்துக்காட்டிற்குப் பின்வரும் பட்டியல் அமைகிறது.

படுகை 1	படுகை 2	s_1^2	s_2^2	$\hat{V}$ $(\bar{y}_{st})$
2, 4	8, 12	2	8	15/36
	8, 16		32	51/36
	12, 16		8	15/36
2, 6	8, 12		8	24/36
	8, 16		32	60/36
	12, 16		8	24/36
4, 6	8, 12		8	15/36
	8, 16		32	51/36
	12, 16		8	15/36

உதாரணமாக, (2, 4, 8, 12) என்ற மாதிரியில்,

$$\hat{V}\,(\bar{y}_{st}) = \frac{1}{36}\left[\,9\left(\frac{3-2}{3}\right)\left(\frac{2}{2}\right)+9\left(\frac{3-2}{3}\right)\left(\frac{8}{2}\right)\right]$$

$$= \frac{15}{36} = 0{\cdot}42$$

இதேபோல் ஏனைய மாதிரிகட்கும் கணிக்கலாம்.

$$\text{மேலும், } E\,[v(\bar{y}_{st})] = \frac{1}{9}\left[\frac{15}{36}+\frac{51}{36}+\dots+\frac{15}{36}\right]$$

$$= \frac{30}{36} = 0{\cdot}83. \quad = V\,(\bar{y}_{st}).$$

∴ முழுமைத்தொகுதி V $(\bar{y}_{st})$-க்குப் பிறழ்ச்சியற்ற மதிப்பீடாகும்.

பயிற்சி

(1) முழுமைத் தொகுதி (2, 3, 3, 4, 12, 14, 15, 15) என்றிருப்பின், அதற்கு σ^2, S^2 மதிப்புகளைக் காண்க. இத் தொகுதியானது இரு படுகைகளாக அமைக்கப்பட்டு (2, 3, 3, 4) என்றும், (12, 14, 15, 15) என்றும் பிரிக்கப்படுகையில், σ_1^2, σ_2^2, S_1^2, S_2^2 இவற்றின் மதிப்புகளைக் கணக்கிட்டு, σ^2 அளவினை நிர்ணயிக்க.

(2) படுகை 1-ல் (2, 12, 15, 3) எனவும். 2-ல் (3, 4, 14, 15) எனவும் எடுத்துக் கொண்டு, மேற் குறிப்பிடப்பெற்ற மதிப்புகளைக் கணக்கிட்டு அறிக.

(3) சென்ற பயிற்சியில், 2 ஆவது கணக்கில், σ_1^2, σ_2^2, S_1^2, S_2^2, σ^2 ஆகியவற்றைத் தீர்மானிக்க.

(4) இரு படுகைகளில்,

I : 2, 4, 6

II : 9, 12, 18, 21

என்ற மதிப்புகள் கொடுக்கப்படின், $n_1 = 2$, $n_2 = 3$ எனும் அளவினதாக எல்லா மாதிரிகளையும் கருதி, $\bar{y}_{st}$-யினுடைய மாதிரி முறை சார்ந்த மாறுபாட்டு அளவினை (sampling variance) கணக்கிடுக. மேலும், அதனுடைய மதிப்பீட்டு அளவையினைத் தீர்மானித்து, $E[v(\bar{y}_{st})] = V(\bar{y}_{st})$ என்று நிறுவுக.

(5) மேற்கண்ட 3ஆவது பயிற்சியில், $V(\bar{y}_{st})$-யினைக் கணக்கிட்டு, அதன் மதிப்பீட்டினையும் பெறுக.

(6) விளைநிலங்களின் சராசரிப் பரப்பு அளவினை, ஒரு பகுதியில் மதிப்பிட்டு அறிதற்காகப் பெறப்பட்ட மாதிரிகள், பின்வருமாறு படுகைப்படுத்தப்பட்டுப் பட்டியலிடப்பட்டுள்ளன :

படுகைகள்	N_h	n_h	y மதிப்புகள் (ஏக்கரில்)
0 — 100 ஏக்கர்	80	5	40, 50, 90, 70, 20, 60.
101 — 200	60	5	140, 150, 140, 130, 180.
201 — 300	40	4	240, 280, 260, 250, 220.
301-க்கு மேல்	20	3	350, 330, 310, 380.

$\bar{y}_{st}$-யினை மதிப்பிட்டு, $V(\bar{y}_{st})$ மற்றும் அதன் மாதிரி மதிப்பீட்டு அளவினையும் தீர்மானிக்க.

படுகை மாதிரிமுறையில் பங்கீடு அல்லது ஒதுக்கீடு (Allocation of the Sample)

முழுமைத் தொகுதியிலிருந்து கருதப்படும் மாதிரியின் அளவான n-மதிப்பையும், துணைமாதிரிகளின் (sub-samples)

அளவுகளாக n_h மதிப்புகளையும் அறியாத நிலையில், $V(\bar{y}_{st})$ மற்றும் $\bar{y}_{st}$ இவற்றின் கணக்கீடுகள் பெறப்படமாட்டா.

எனவே, மாதிரி அளவு n-ஐ தீர்மானித்து இந்த n மதிப்பை h படுகைகளில் தக்கபடி பங்கிட்டு (allocate) அல்லது தகுந்தபடி அளவுகளை நிர்ணயித்து, n_h மதிப்பைத் தீர்மானிப்பது கருதப் படும். அதாவது, 'எந்த முறையில் n மதிப்பினை, படுகைகளுள் பங்கிட்டு மாதிரிகளைத் தெரிவுசெய்தால் சிறந்த முடிவுகள் பெறப்படும்?' என்பதே பிரச்சினை. அதாவது, ஒரு குறிக் கோளுடன், பல்வேறு படுகைகளிலிருந்து தெரிவு செய்யப்படும் மாதிரிகளின் அளவுகளைத் தீர்மானிப்பது பங்கீடுபற்றிய அல்லது ஒதுக்கீடுபற்றிய பிரச்சினை எனப்படும்.

ஒரு சிறந்த மாதிரி அளவுப் பங்கீட்டு முறையானது (allocation of sample),

(i) கொடுக்கப்பட்ட செலவிற்குத் தக்கபடி மாதிரிமுறை சார்ந்த மாறுபாட்டினை (sampling variance) குறைக்கும்படி யாகவோ, அல்லது

(ii) கொடுக்கப்பட்ட மாதிரிமுறை சார்ந்த மாறுபாட்டு அளவிற்கேற்ப, செலவினைக் (cost) குறைக்கும்வண்ணமோ அமைதல் வேண்டும்.

மாதிரிமுறையில் செலவானது வடிவமைப்புபற்றியோ, கூறுகளை அமைப்பதிலோ, விவரங்களைச் சேகரிப்பதிலோ அல்லது சில மாருச் செலவுகளாகவோ அமைந்திருக்கலாம். எளியதொரு செலவுச் சார்பானது (cost function),

$$C = C_0 + C_h \cdot n_h$$

என்று ஒரு படுகைக்குக் (stratum) கருதப்படலாம். இதில் C_0 என்பது மாருச் செலவுகளையும் (overhead costs), C_h என்பது ஒரு படுகையிலிருந்து இன்னொரு படுகைக்கு மாறும் இயல்புடைய (variable), ஒருகூறு பற்றிய (சராசரி) செலவு. இது, மாதிரியின் அளவைப் பொறுத்து (n_h) மாறுபடும்.

உத்தமப் பங்கீடு (Optimum Allocation)

வெவ்வேறு படுகைகளிலிருந்து மாதிரிக் கூறுகள் எடுக்கும் செலவுகள் மாறும் தன்மையில் இருக்கலாம். உதாரணமாக, குடும்பங்களை நகர மற்றும் கிராமிய அடிப்படையில் படுகைப் படுத்தி மாதிரியிடும்போது, அவற்றிற்கு ஆகும் செலவுகள் மாறக்கூடும். செலவுச் சார்பு.

$$C = C_0 + \sum_{h=1}^{L} C_h n_h$$

என்று கொள்வோம். அல்லது

$$(C - C_0) = \sum_{h=1}^{L} (C_h . n_h) \text{ என்று கருதுக.}$$

இந்தக் கொடுக்கப்படும் செலவுத் திட்ட அடிப்படையில், n அளவுள்ள ஒரு மாதிரியைத் தெரிவுசெய்து, மேற்கண்ட உதாரணப்படி, அதனை n_1, n_2 இவற்றில் தக்கபடி பங்கீடு செய்வது, $V(\bar{y}_{st})$ மதிப்பை மிகக் குறைந்த அளவர்க்கும் வண்ணம் அமைவதானது 'உத்தமப் பங்கீட்டுமுறை' (Optimum allocation) எனப்படும்.

தேற்றம் 5

படுகை மாதிரிமுறையில், செலவுச் சார்பினை $C = C_0 + \sum_{h=1}^{L} C_h . n_h$ எனக் கொண்டு, n அளவைப் பங்கீடு செய்கையில், n_h எனும் அளவானது $\frac{N_h . S_h}{\sqrt{C_h}}$ என்னும் மதிப்புடன் வீதத் தொடர்புடையதாய் (proportional) அமையும் நிலையில், $\bar{y}_{st}$ மதிப்பீட்டின் மாதிரி முறைசார் மாறுபாடானது மீச்சிறு (minimum) அளவினதாகும்.

நிருபணம்

கருதப்படும் பிரச்சினையாவது:

$$(C - C_0) = C_1 . n_1 + C_2 . n_2 \text{ ....} + C_L . n_L$$

என்ற நிபந்தனைக்கேற்றவண்ணம்,

$$V(\bar{y}_{st}) = \sum_{h=1}^{L} \frac{W_h^2 . S_h^2}{n_h} (1 - f_h)$$

$$= \sum_{h=1}^{L} \frac{W_h^2 . S_h^2}{n_h} - \sum_{h=1}^{L} \frac{W_h^2 . S_h^2}{N_h}$$

என்ற மாறுபாட்டை மீச்சிறு அளவினதாக்கல் வேண்டும்.

நுண்கணிதமுறை 'இலக்ராஞ்' சின் (Lagrange s) பெருக்கு எண்களைப் (Multipliers) பயன்படுத்தி, இப் பிரச்சினையைக் கருதுவோம்.

n_h-ஐயும், பெருக்கு எண் λ வினையும்,

$$V(\bar{y}_{st}) + \lambda \left[\sum_{h=1}^{L} C_h \cdot n_h - C + C_0 \right]$$

என்பதனைச் சிறு அளவினதாக்கும் வகையில் தெரிவு செய்தல் வேண்டும்.

அதற்கு

$$\Sigma \frac{w_h^2 S_h^2}{n_h} - \Sigma \frac{w_h^2 \cdot S_h^2}{N_h} + \lambda \left(C_1 n_1 + \ldots\ldots + C_L n_L - C + C_0\right)$$

என்பதை, n_h-ஐத் தொடர்புபடுத்தி வகையிடும்போது (differentiate) கிடைக்கும் சமன்பாடுகள்.

$$- \frac{w_h^2 \cdot S_h^2}{n_h^2} + \lambda \cdot C_h = 0$$

$$h = 1, 2 \ldots\ldots L$$

ஆகும்.

அதாவது,

$$n_h \cdot \sqrt{\lambda} = \frac{w_h S_h}{\sqrt{C_h}} ; \quad h = 1, 2 \ldots\ldots L. \qquad \ldots\ldots (1)$$

இதனை, எல்லாப் படுகைகளையும் கருதிச் சராசரியிடும்போது,

$$n \cdot \sqrt{\lambda} = \sum_{h=1}^{L} \frac{w_h \cdot S_h}{\sqrt{C_h}} \qquad \ldots\ldots (2)$$

ன்று பெறப்படும்.

(1), (2) ஆகியவற்றின் விகித அளவு

$$\frac{n_h}{n} = \frac{\dfrac{w_h \cdot S_h}{\sqrt{C_h}}}{\displaystyle\sum_{h=1}^{L} \frac{w_h \cdot S_h}{\sqrt{C_h}}} = \frac{\dfrac{N_h\, S_h}{C_h}}{\displaystyle\sum_{h=1}^{L} \frac{N_h\, S_h}{\sqrt{C_h}}}$$

எனப் பெறப்படுகிறது. இதனின்று,

$$n_h = \frac{\dfrac{N_h \cdot S_h}{\sqrt{C_h}}}{\displaystyle\sum_{h=1}^{L} \frac{N_h \cdot S_h}{\sqrt{C_h}}} \cdot n$$

எனப் பெறலாம்.

இத் தேற்றத்தின்மூலம், ஒரு படுகையானது (stratum) பெரிய அளவினதாக (large size) இருப்பின், உள்ளீட்டு மாறு பாட்டை அதிக அளவில் பெற்றிருப்பின், மேலும் அதனின்று குறைந்த செலவில் மாதிரியினைத் தெரிவு செய்ய முடிந்திடின்,

ஒரு படுகையிலிருந்து பெரிய அளவுள்ள மாதிரியினை எடுப்பது நன்று எனப் புலப்படும்.

மேற்கண்ட முடிவில், n_h-ன் மதிப்பினை n-ன் மூலமாகக் கண்டறிந்தோம். ஆனால், n-ன் மதிப்பு என்னவெனத் தெரிந்திராத நிலையில், தீர்வு (solution) முற்றுப் பெறுவதில்லை.

முழு அளவிலான தீர்வானது, குறிக்கப்பட்ட செலவிற் கேற்ப அல்லது $\bar{y}_{st}$ குறிக்கப்பட்ட மாறுபாட்டிற்கேற்ப மாதிரி யானது தெரிவு செய்யப்படுகிறதா என்பதைப் பொறுத்தது. செலவினை நிலையானதாக அல்லது மாறுததாகக் (fixed) கருது வோம். n_h-ன் உத்தம மதிப்பினை (Optimum value) செலவுச் சார்பில் பிரதியிடுவோம். அதன்மூலம் n மதிப்புக்குத் தீர்வு காணலாம்.

அதாவது,

$$C = C_0 + \sum_{h=1}^{L} C_h\, n_h$$

அல்லது

$$(C - C_0) = \sum_{h=1}^{L} C_h n_h$$

$$= \sum_{h=1}^{L} \frac{\dfrac{C_h \,.\, n \,.\, N_h \,.\, S_h}{\sqrt{C_h}}}{\sum_h \dfrac{N_h \,.\, S_h}{\sqrt{C_h}}}$$

என்பதிலிருந்து,

$$n = \frac{(C - C_0) \,.\, \sum_h \dfrac{N_h S_h}{\sqrt{C_h}}}{\sum_h N_h \,.\, S_h \,.\, \sqrt{C_h}}$$

என்று பெறலாம்.

அல்லது, $\bar{y}_{st}$-ன் மாறுபாடான $V(\bar{y}_{st})$, அதாவது V ஆனது நிலையானதெனக் கொள்கையில், n_h-ன் உத்தம மதிப்பைப் பிரதியிட,

$$V = V(\bar{y}_{st}) = \sum_h \frac{w_h^2 \,.\, S_h^2}{n_h} - \sum_h \frac{w_h S_h^2}{N}$$

என்பது

$$V + \frac{1}{N} \sum_h w_h^2 S_h^2 = \sum_{h=1}^{L} \frac{w_h^2 \,.\, S_h^2 \,.\, \left(\dfrac{\Sigma\, w_h S_h}{\sqrt{C_h}}\right)}{\dfrac{n.\, w_h S_h}{\sqrt{C_h}}}$$

$$= \sum_h \frac{w_h S_h \left(\sum_h \dfrac{w_h S_h}{\sqrt{C_h}}\right)}{\dfrac{n}{\sqrt{C_h}}}$$

என்றாக, அதன்மூலம்

$$n = \frac{\left[\sum_h w_h S_h \cdot \sqrt{C_h}\right] \left[\sum_h \frac{w_h S_h}{\sqrt{C_h}}\right]}{V + \frac{1}{N} \sum_h w_h S_h^2}$$

$$\left[w_h = \frac{N_h}{N}\right]$$

எனப் பெறப்படும்.

முக்கியமான ஒரு குறிப்பிட்ட (special) நிலை

செலவானது எல்லாப் படுகைகட்கும் ஒன்றுனதாக (same) இருந்தால், அதாவது $C = C_h'$ எனில், செலவு நிபந்தனை யானது, மொத்த மாதிரி அளவினை (n) தீர்மானிக்கும்பொருட்டு அமையும். இந் நிலையில் செலவுச் சார்பானது (cost function),

$$C = C_0 + C'_n \quad \text{என்றுக,}$$

மாதிரி அளவு (n) நிலையானதாகக் கருதப்பட்டு, உத்தமப் பங்கீட்டுப் பிரச்சினை மாறுகிறது. பின்வரும் முடிவினைக் கருதுவோம்.

தேற்றம் 6

படுகையாக்கும் மாதிரி முறையில், மொத்த மாதிரி அளவு (total sample size) n ஆனது நிலையானதாகக் (fixed) கொள்ளப் படின், படுகை மாதிரி அளவு

$$n_h = n \cdot \frac{w_h S_h}{\Sigma w_h S_h} = n \cdot \frac{N_h \cdot S_h}{\Sigma N_h S_h}$$

என்றிருப்பின், $V(\bar{y}_{st})$-ன் மதிப்பு மீச்சிறு அல்லது சிறும அளவினதாகும்.

குறிப்பு : இத்தகைய, மாறுத அல்லது நிலையான மாதிரி அளவுடன் பெறப்படும் பங்கீட்டு முறையானது, 'நேமன் பங்கீடு' (Neyman Allocation) எனப்படும் (பின்னுல் இதை விரிவாகக் கருதுவோம்).

n மதிப்பினை மாறுததாகக் (நிலையாக) கொண்டு, சிறும அளவு (மாதிரி முறை சார்ந்த) மாறுபாட்டினை,

$$V(\bar{y}_{st}) = \sum_h \frac{w_h^2 \cdot s_h^2}{n_h} - \sum_h \frac{w_h \cdot s_h^2}{N}$$

என்ற நியதியில் n_h-க்கு மேற்கண்ட தேற்றத்தில் கொடுக்கப்பட்ட மதிப்பைப் பிரதியிட,

$$\underset{(\text{சிறும அளவு})}{(V\bar{y}_{st})} = \sum_h \frac{w_h^2 \cdot S_h^2}{n \cdot w_h \, S_h} \cdot (\Sigma \, w_h \, S_h) - \sum_N \frac{w_h \, S_h^2}{N}$$

$$= \frac{\left[\sum_h w_h \, S_h\right]^2}{n} - \frac{\sum_h w_h \, S_h^2}{N}$$

வலப்புறம் இரண்டாவது உறுப்பு (term) '*fpc*' ஆகும்.

உத்தமப் பங்கீட்டு (optimum allocation) முறையில் பெறக் கூடிய $\bar{y}_{st}$ மதிப்பீட்டின் முறைசார்ந்த மாறுபாட்டு அளவினை (sampling variance) V_{opt} எனக் குறிப்பிட்டு, அதை

$$V(\bar{y}_{opt}) = V_{opt}$$

$$= \frac{1}{N^2} \cdot \frac{1}{n} \cdot \left(\sum_h N_h S_h \cdot \sqrt{C_h}\right)\left(\sum_h \frac{N_h \, S_h}{\sqrt{C_h}}\right)$$

$$- \frac{1}{N^2} \sum_h N_h S_h^2$$

என்று நடைமுறையில் கணக்கிட்டு அறியலாம்.

படுகைப்படுத்திய, எளிய சமவாய்ப்பு மாதிரிமுறையினில் (Stratified, SRS), உத்தமப் பங்கீட்டு முறையினைச் சிறந்த முறையில் பயன்படுத்துவதற்கு, முந்திய ஆய்வுகளில் மதிப்பிடப்பட்ட, படுகைகளின் மாறுபாட்டு (stratum variances) மதிப்பீடுகளைக் கருதுவதானது, மதிப்பீட்டுத் திட்பத்தினை அதிகரிக்கும் என்பது கண்டறியப்பட்ட உண்மையாகும்.

செலவுகளின் விகித (ratio) அல்லது மாறுபாடுகளின் விகித அளவு பல்வேறு படுகைகட்கும் அறியப்பட்டிருப்பின், n_h மதிப்பைக் கண்டறிவது எளிதாக அமையும். அவை அறியப் பெறாத நிலையில், மதிப்பீடுகள் (estimates) பயன்படுத்தப்பெறும். பல்வேறு படுகைகளிடையில் மாறுபாட்டு அளவு அல்லது செலவு (s_h அல்லது C_h) ஏறக்குறைய ஒரேமாதிரி அமையும் நிலையில், $s_1 = s_2$ எனவும், $C_1 = C_2$ எனவும் கொண்டு,

$$n_h = \frac{N_h}{N} \cdot n$$

என்பதாக மதிப்பிடப்பெறுகையில், இது விகிதசம அல்லது வீதப் பங்கீட்டு நிலைக்கு (proportional allocation) சமமாக அமைவதைக் குறிப்பிடலாம். படுகைகள் ஒன்றோடொன்று அதிக அளவில் வேறுபட்டு நிற்கும் நிலையில், உத்தமப் பங்கீடு நல்ல பயன் தருவதாகும்.

எடுத்துக்காட்டு 1 :

உத்தமப் பங்கீட்டு முறையில் (optimum allocation) படுகை மாதிரி அளவுகளை நிர்ணயித்தல் :

படுகை 1 : 2, 2, 4, 6, 6

படுகை 2 : 8, 8, 12, 16, 16.

மேற் கொடுக்கப்பட்ட முழுமைத் தொகுதிக்கு, உத்தமப் பங்கீட்டு முறையில், மாதிரி அளவான $n = 4$ என்பதைப் பங்கீடு செய்து, முழுமைத் தொகுதி சராசரிக்கும், மாதிரிமுறை சார்ந்த மாறுபாட்டுக்கும் (Sampling variance) மாதிரி மதிப்பீட்டு அளவைத் (estimators) தீர்மானிக்கவும். $C_1 = 1$; $C_2 = 4$ எனக் கொள்க.

(i) n_h — மதிப்புகளைக் காணுதல்.

$$S_1^2 = \frac{1}{4} \Sigma (y_{1i} - \bar{Y}_1)^2$$

$$= \frac{1}{4} [(-2)^2 + (-2)^2 + 0^2 + (2)^2 + 2^2] = 4$$

இதுபோல், $S_2^2 = 16$; எனவே, $S_1 = 2$; $S_2 = 4$.

∴	N_h	S_h	S_h^2	C_h	$\sqrt{C_h}$	N_hS_h	$\frac{(N_hS_h)}{\sqrt{C_h}}$
I	5	2	4	1	1	10	10
II	5	4	16	4	2	20	10
							20

வரைமுறைப்படி,

$$n_1 = \frac{\frac{N_1 S_1}{\sqrt{C_1}}}{\Sigma \frac{(N_h S_h)}{\sqrt{C_h}}} . \quad n = \frac{10}{20} . \; (4) = 2 \text{ எனவும்}$$

$$n_2 = \frac{\frac{N_2 S_2}{\sqrt{C_2}}}{\Sigma \frac{N_h S_h}{\sqrt{C_h}}} = \frac{10}{20} (4) = 2$$

எனவும் பெறப்படுகின்றன.

(ii) $\bar{y}_{st}$ யினைக் காணல் :

ராண்டம் பட்டியலிலிருந்து $n_1 = 2$; $n_2 = 2$ அளவு மாதிரிக் கூறுகளைப் படுகை I லிருந்தும், படுகை II லிருந்தும் முறையே தெரிவு செய்து, இரு மாதிரிகளை (4, 6), (8, 16) என்று y-ன் மதிப்புகள் அமையுமிடத்து, $y_1 = 10$, என்றும் $y_2 = 24$ என்றும் பெற்று, $\bar{y}_1 = 5$ எனவும் $\bar{y}_2 = 12$ எனவும் மதிப்பிடப்படும்.

$$\therefore \quad \bar{y}_{st} = \frac{N_1 \bar{y}_1 + N_2 \bar{y}_2}{N_1 + N_2} = \frac{25 + 60}{10} = \frac{85}{10} = 8 \cdot 5.$$

$\bar{y}_{st}$யின் மாதிரிமுறை சார் மாறுபாடு (Sampling variance): அதாவது, V_{opt} மதிப்பைக் காணல்.

படுகை	$(N_h S_h) \sqrt{C_h}$	$N_h S_h^2$	$(N_h S_h)^2$	$\frac{(N_h S_h)^2}{n_h}$
I	10	20	100	50
II	40	80	400	200
	50	100		250

$$\therefore \quad V_{opt} = \frac{1}{N^2} . \frac{1}{n} \left[\sum_h N_h S_h \cdot \sqrt{C_h} \right] \left[\sum_h \frac{N_h S_h}{\sqrt{C_h}} \right] - \frac{1}{N^2} \sum_h N_h S_h^2 .$$

$$= \frac{1}{10^2} \cdot \frac{1}{4} \cdot (50)\ (20) - \frac{1}{10^2}\ (100)$$

$= 2{\cdot}5 - 1{\cdot}0 = 1{\cdot}5$ ஆகும்.

மற்ருெரு முறை :

$$V_{(\bar{y}st)} = \frac{1}{N^2} \Sigma \frac{N_h^2 S_h^2}{n_h} - \frac{1}{N^2} \Sigma N_h S_h^2$$

என்பதில், n_h மதிப்புகளை ($n_1 = 2$; $n_2 = 2$) பிரதியிட,

$$V_{opt} = \frac{1}{10^2} \cdot (250) - \frac{1}{10^2} \cdot (100)$$

$= 2{\cdot}5 - 1{\cdot}0 = 1{\cdot}5$

என்று பெறலாம்.

பயிற்சி

$V(\bar{y}_{opt})$ அதாவது V_{opt}-ன் மாதிரி மதிப்பீட்டைக் காண்க. (S_h க்குப் பதில் s_h மதிப்பைக் கருதுக).

வீதப் பங்கீடு (Proportional Allocation)

ஒவ்வொரு துணைத்தொகுதியின் (sub-population) அளவினை (N_h) தவிர வேறு எந்த விவரமும் அறியப் பெருதநிலையில், ஒரு மாதிரி அளவான n-ஐப் பங்கிடும் முறையானது, N_h-உடன், வீதத் தொடர்புள்ளதாய் அமையும். அதாவது n_h ஆனது அப் படுகை அளவான N_h உடன், வீதத் தொடர்புள்ளதாய் இருக்கும் (proportional). உதாரணமாக, $N = 400$ என்ற முழுமைத் தொகுதியிலிருந்து, $n = 40$ அளவுள்ள (size) மாதிரியினைத் தெரிந்தெடுக்கையில், மாதிரி எடுத்தற் பின்னமானது (sampling fraction) $\frac{n}{N} = \frac{40}{400} = 0{\cdot}1$ ஆவதால், பங்கீட்டு முறையானது (allocation) $\frac{n_1}{N_1} = \frac{n_2}{N_2} = \ldots\ldots = \frac{n}{N} = 0{\cdot}1$ என்ற நியதியினைப் பின்பற்றி அமையும்.

பொதுவாக, $n_h = n.\ W_h$ என்றிருக்கவேண்டும். $\left(W_h = \frac{N_h}{N}\right)$ இம் முறையானது, கணக்கீட்டு எளிமை கருதியும், மாதிரிமுறை

சார்ந்த மாறுபாட்டினை (sampling error) பொருட்படுத்தாத நிலையிலும், கையாளப்படுகிறது.

கொடுக்கப்பட்ட செலவுக்கேற்ப, மாதிரி அளவினை (n), இம் முறையில், $n_h = n.W_h$ என்று செலவுச் சார்பில் பிரதியிட்டு (substitute), n-ன் மதிப்பிற்கு தீர்வு காண்பதின் மூலமாகப் பெறலாம். அத்தீர்வு

$$n = \frac{(C - Co)}{\sum_h W_h.C_h} \text{ ஆகும்.}$$

எடுத்துக்காட்டு :

வீதப் பங்கீட்டின் அடிப்படையில் மாதிரித் தெரிவு : இரு படுகைகளை (strata) கருதுக :

படுகை 1 : 2, 4, 6

படுகை 2 : 8. 12, 16.

இதில், $N_1 = N_2 = 3$; $N = 6$.

$n = 4$ என்ற அளவினதான மாதிரியினைத் தெரிவுசெய்ய, வீதப் பங்கீட்டு முறைப்படி,

$$\frac{n_1}{N_1} = \frac{n_2}{N_2} = \frac{n}{N} = \frac{4}{6} = \frac{2}{3}.$$

எளிய சமவாய்ப்பு முறைப்படி, ஒவ்வொரு கூறினுக்கும் தெரிவு செய்யப்படும் வாய்ப்பானது, $\frac{4}{6} = \frac{2}{3}$ என்ற அளவில் அமையும்.

படுகை மாதிரி அளவுகள், $n_h = \frac{N_h}{N} \times n$ என்றபடி,

$$n_1 = \frac{N_1}{N} n = \frac{2}{3} 3 = 2$$

$$n_2 = \frac{N_2}{N} n = \frac{2}{3} 3 = 2$$

என்று பெறப்படும்.

$\bar{y}_{st}$-யின் கணக்கீடு

வரையறைப்படி, $\bar{y}_{st} = \dfrac{\Sigma N_h . \bar{y}_h}{\Sigma N_h}$

இதில், $N_h = \dfrac{n_h}{f} \left(\because \dfrac{n}{N} = f \right)$ என்று பிரதியிட்டு,

$$\bar{y}_{st} = \frac{\Sigma \dfrac{n_h}{f} \left(\dfrac{\Sigma y_{hi}}{n_h} \right)}{\Sigma \dfrac{n_h}{f}}$$

$$= \Sigma \frac{n_h . \bar{y}_h}{n} = \frac{1}{n} \sum_h \sum_i y_{hi}$$

அதாவது, $\bar{y}_{st}$ என்பது n- அளவுடைய மாதிரியின் சராசரி மதிப்பாகும். இது-$\bar{Y}$ன் பிறழ்ச்சியற்ற மதிப்பீடு.

உதாரணத்தில், (2, 4), (8, 16) என்ற மாதிரிகளைக் கருதுவோம்.

இவற்றில், $y_{11} = 2$; $y_{12} = 8$

$y_{21} = 4$; $y_{22} = 16$

$$\bar{y}_{st} = \frac{2 + 8 + 4 + 16}{4} = \frac{30}{4} = 7{\cdot}5$$

ஆகும்.

$V (\bar{y}_{st})$ -ன் மதிப்பு

ஏற்கெனவே கண்டபடி

$$V (\bar{y}_{st}) = \frac{1}{N^2} \Sigma N_h^2 \frac{N_h - n_h}{N_h} \frac{S_h^2}{n_h}$$

இதில், $n_h = \left(\dfrac{N_h}{N} n \right)$ என்று பிரதியிட்டு,

$$V_{prop} = V (\bar{y}_{prop}) = \frac{N-n}{N} \sum_L \frac{N_h}{N} \frac{S_h^2}{n}$$

என்று பெறலாம். $(f \neq 1)$

$f = 1$ என்ற நிலையில், $V_{prop} = \displaystyle\sum_L \frac{N_h}{N} \frac{S_h^2}{n}$

இதில், S_h^2க்கு அதன் பிறழ்ச்சியற்ற மதிப்பீடான s_h^2-ஐ பிரதியிட்டு, V_{prop}-ன் பிறழ்ச்சியற்ற மதிப்பீட்டை,

$$\hat{V}_{prop} = \frac{N-n}{N} \sum_{L} \frac{N_h}{N} \frac{s_h^2}{n} \quad (fpc \neq 1)$$

$$= \sum_{L} \frac{N_h}{N} \frac{s_h^2}{N} \quad (fpc = 1)$$

என்று பெறலாம்.

கருதப்பட்ட உதாரணத்தில், முதல் படுகையிலிருந்து (2, 4) என்ற மாதிரியும், இரண்டாவதில் (8, 16) என்பதும் எடுக்கப்பட்டன.

$$\therefore s_1^2 = (2-3)^2 + (4-3)^2 = 2$$

$$s_2^2 = (8-12)^2 + (16-12)^2 = 32$$

∴ மதிப்பிட்ட மாறுபாடு,

$$\hat{V}_{prop} = \frac{N-n}{N} \sum \frac{N_h}{N} \cdot \frac{s_h^2}{n}$$

$$= \frac{(6-4)}{6} \left[\left(\frac{3}{6} \cdot \frac{2}{4} \right) + \left(\frac{3}{6} \cdot \frac{32}{4} \right) \right]$$

$$= \frac{2}{6} \cdot \left[\frac{51}{12} \right] = \frac{17}{12} = 1{\cdot}42$$

ஆகிறது.

பயிற்சி

(1) படுகை I : 2, 4, 6

படுகை II : 8, 12, 16

இவற்றிருந்து எடுக்கப்படும் மற்ற மாதிரிகளைக் கருதி, அவற்றின் அடிப்படையில், $V(\bar{y}_{prop})$ மதிப்பைக் கணக்கிடுக.

(2) ஆண்களும், பெண்களும் படித்துமுடிக்கும் புத்தகங்கள் படுகைகளில் தரப்பட்டுள்ளன.

படுகை 1 (ஆண்) : 20 25 30 26 24 5

படுகை 2 (பெண்) : 15 20 10 3

வீதப் பங்கீட்டு முறையில், $n = 5$ என்ற சமவாய்ப்பு மாதிரியினைத் தெரிவு செய்து, சராசரியினை மதிப்பிடுக.

(ii) (20, 25, 24) மற்றும் (20, 10) என்ற மாதிரிகளைக் கொண்டு, V_{prop}-ன் மதிப்பீட்டைப் பெறுக.

(3) 500 விற்பனையகங்களை (stoses), மூன்று படுகைகளாகப் பிரித்து, பின்வருமாறு கொடுக்கப்படுவதன் அடிப்படையில்

(i) சராசரி விற்பனை அளவு

(ii) அதன் மாதிரிமுறைசார் மாறுபாட்டு அளவு இவற்றை மதிப்பிடுக. மாதிரி அளவு $n = 40$ எனக் கொள்க.

படுகை	N_h	$\bar{x}_h$	s_h^2
சிறியது	300	40	1
நடுத்தர அளவு	150	120	4
பெரியது	50	240	16
	500		

$\bar{x}_h$, s_h^2 இரண்டும், ' ரூபாய் ஆயிரம் ' எனும் அலகுப்படி அமைவதாகக் கொள்க.

(4) இரு படுகைத் தொகுதிகள்.

I : 5, 9, 4, 5, 2, 8, 4, 5, 2

II : 13, 16, 18, 16, 18, 17, 19, 17, 16, 11, 18, 14, 13, 15, 12 என்று கொடுக்கப்படுகையில், $n = 8$ அளவில், வீதப் பங்கீட்டு முறையில் மாதிரியைத் தெரிவு செய்து, அந்த மாதிரி அடிப்படையில் $\bar{y}_{st}$ மற்றும் V_{prop} இவற்றின் மதிப்புகளைக் காண்க.

(5) 1000 விற்பனைக் கூடங்களின் விற்பனை அளவுகள் பணி புரிவோர் அடிப்படையில் இரு பிரிவுகளாகப் பகுக்கப்படுகின்றன.

–	N_h	$\bar{y}_h$	s_h^2
1-ம் பிரிவு	600	5	2
2-ம் பிரிவு	300	15	8
3-ம் பிரிவு	100	30	32

(y ஆனது 1000 ரு. என்ற அலகின் அடிப்படையில் கருதப் படுகிறது).

வீதப்பங்கீடு முறையில், மொத்த விற்பனை, மாதிரிசார் மாறுபாடு இவற்றைக் கணக்கிடுக. $n = 80$ எனக் கருதுக.

மேலும், n_h, $\hat{V}(\bar{y}_{prop})$ இவற்றை கணக்கிடுக.

(6) பின்வரும் பட்டியலைப் பயன்படுத்தி,

(i) உத்தமப்பங்கீடு முறையில் n-அளவினைப் பங்கிடுக.

(ii) $\bar{y}_{st}$, $\hat{V}_{opt}$ இவற்றின் மதிப்புகளைக் கணக்கிடுக. தரப்பிழையினையும் தீர்மானிக்க.

படுகை	N_h	$\bar{x}_h$	s_h	c_h
1	80	50	12	9
2	160	30	8	4
3	260	10	4	1

n—என்பதை 100 எனக் கொள்க.

(7) படுகை 1 : 2, 4, 4, 6, 6, 8

2 : 8, 8, 12, 16, 16, 18

3 : 16, 16, 18, 22, 22, 26

என்ற தொகுதியினின்று, $n = 10$ என்ற அளவை, உத்தமப் பங்கீட்டு முறையில் பங்கிடுக. $V(\bar{y}_{opt})$ என்ன ?

($c_1 =$ ரூ. 2 ; $c_2 =$ ரூ. 4 ; $c_3 =$ ரூ. 5 என்று கொள்க).

(8) மேற்கண்ட கணக்கில், சில சமவாய்ப்பு மாதிரிகளைக் கண்டறிந்த அளவுப்படி (n_h) தெரிவு செய்து, $\bar{Y}$-ன் மதிப்பீட்டைப்பெறுக.

(9)

N_h	s_h	$\bar{x}_h$
80	12	50
160	8	30
260	4	10

என்பதில், $n = 100$ எனும் மாதிரி அளவினை, உத்தமப் பங்கீட்டு முறைப்படி பங்கிடுக. ($c_1 = 9$; $c_2 = 4$; $c_3 = 1$ எனக் கொள்க). $\bar{y}$ மற்றும் $s(\bar{y}_{opt})$ இவற்றையும் கணக்கிடுக.

(10) $V(y_{opt})$-லிருந்து, $V(\hat{Y}_{opt})$-ன் மதிப்பைக் காண்க.

‘நேமன்’ பங்கீடு (Neyman Allocation) (1934)

செலவு C_h ஆனது எல்லாப் படுகைகட்கும் சமமானதாகக் கருதப்படும் நிலைகளில், $C_h = C_f$ எனக் கொண்டு,

செலவுச் சார்பை $C = Co + C_f \Sigma n_h = Co + C_f . n$ எனக் கருதி,

$$n = \frac{C - Co}{C_f}$$

எனப் பெற்று, n-என்ற மாதிரி அளவானது நிலையானதாக ஆகிறது. உத்தமப் பங்கீட்டு முறை முடிவுப்படி, $C_h = C_f$ = மாறிலி (constant) என்ற பிரதியீட்டுடன்

$$n_h = \frac{N_h S_h}{\Sigma N_h S_h} . \ n \qquad \ldots\ldots \quad (1)$$

ஆகிறது. $\therefore$ $N_h.S_h$ என்ற மதிப்போடு வீதத் தொடர்புடையதாக n-ன் பங்கீடு அமைகிறது.

பெரியதொரு அளவினதாகவும், பல்வேறு இயல்பினதாகவும் (heterogeneous) உள்ள படுகைகளினின்று, பெரிய (large size) அளவில் மாதிரிகளை எடுத்தல் வலியுறுத்தப்படுகிறது.

(1)-ன் அடிப்படையில் தீர்மானிக்கப்பட்ட n_h அளவினதான துணைக் கூறுகள் (sub-samples) எடுத்து, $\bar{y}_h$-ன் மதிப்புகளை கணக்கிட்டுப் பெற்று, அதன் மூலமாக

$$\bar{y}_{st} = \frac{\Sigma N_h . \bar{y}_h}{N}$$

பெறப்படும். மேலும் $E(\bar{y}_{st}) = \bar{Y}$.

$\bar{y}_{st}$ யின் மாதிரிசார் மாறுபாடானது

$$V_{Ney} = V(\bar{y}_{Ney}) = \frac{1}{N^2} . \frac{\left(\sum_h N_h \ s_h\right)^2}{n} - \frac{\sum N_h \ s_h^2}{N^2}$$

என்று மதிப்பிடப்பெறும்.

எடுத்துக்காட்டு : கொடுக்கப்படும் பட்டியல் :

படுகை	N_h	s_h	$\bar{y}_h$
1	80	12	50
2	160	8	30
3	260	4	0
	500.		

$n = 100$ என்று கொண்டு, நேமன் முறைப் பங்கீட்டின்படி, $\bar{y}_{st}$, $\hat{V}(\bar{y}_{st})$ இவற்றை மதிப்பிடுக.

முதலில் 'n_h' தீர்மானிக்கப்படுதல் :

வரையறை :

$$n_h = \frac{N_h\, s_h}{\Sigma\, N_h \cdot s_h} \cdot n.$$

இதன்படி,

$$n_1 = \frac{960}{3280} \times 100 \doteq 29$$

$$n_2 = \frac{1280}{3280} \times 100 \doteq 39$$

$$n_3 = \frac{1040}{3280} \times 100 \doteq 32$$

	$N_h\, s_h$	$N_h \cdot \bar{y}_h$
1.	960	4000
2.	1280	4800
3.	1040	2600
	3280	11400

என்று பெறப்படும்.

மேலும்,

$$\bar{y}_{st} = \frac{\Sigma\, N_h \cdot \bar{y}_h}{N}$$

$$= \frac{11400}{500} = 22{\cdot}8$$

$\hat{V}(\bar{y}_{Ney})$ கணக்கிடல் :

n_h	$(N_h \cdot s_h)^2$	$\frac{(N_h \cdot s_h)^2}{n_h}$	$\frac{(N_h - n_h)}{N_h}$	
(1)	(2)	(3)	(4)	(3) × (4)
29	$(960)^2$	$\frac{(960)^2}{29}$	$\frac{51}{80}$	20,000
39	$(1280)^2$	$\frac{(1280)^2}{39}$	$\frac{121}{160}$	31,000
32	$(1040)^2$	$\frac{(1040)^2}{32}$	$\frac{228}{260}$	30,000
100				81,000

$$\therefore \quad \hat{V}_{Ney} = \hat{V} \quad (\hat{y}_{Ney})$$

$$= \frac{1}{N^2} \sum \frac{N_h - n_h}{N_h} \cdot \frac{(N_h \cdot s_h)^2}{n_h}$$

$$\therefore \quad \hat{V}_{Ney} = \frac{1}{(500)^2} \cdot (81000) = 0{\cdot}32$$

இதன் வர்க்கமூலம்

$s(\hat{y}_{Ney}) = 0{\cdot}57$ ஆகும்.

எடுத்துக்காட்டு

நேமன் பங்கீடு:

படுகை	N_h	s_h	$N_h \cdot s_h$	n_h	$\bar{y}_h$	$N_h \cdot \bar{y}_h$
I	80	12	960	29	50	4000
II	160	8	1280	39	30	4800
III	260	4	1040	32	10	2600
			3280			

(s_h-ன் மதிப்புகள் ஏற்கெனவே கண்டறியப்பட்டவை.)

$n = 100$ எனக் கொண்டு, பங்கீடு செய்தலைக் கருதுவோம்.

நியதிப்படி, $n_h = \frac{N_h\, s_h}{\Sigma N_h\, s_h} \cdot n$ ஆகும்.

$$\therefore \quad n_1 = \frac{N_1\, s_1}{\Sigma N_h\, s_h} \cdot n$$

$$= \frac{960}{3280} \cdot 100 \doteqdot 29.$$

$$n_2 = \frac{1280}{3280} \cdot 100 \doteqdot$$

$$n_3 = \frac{1040}{3280} \cdot 100 \doteqdot 32$$

என்று மதிப்பிடப்படுகின்றன.

$\bar{y}_{st}$ **காணல் :**

$\bar{y}_h$-க்குக் கொடுக்கப்பட்ட மதிப்புகள் 50, 30, 10. இதன் மூலம்,

$$\bar{y}_{st} = \frac{\Sigma N_h \cdot \bar{y}_h}{N} = \frac{11400}{500} = 22{\cdot}8.$$

$\hat{V}$ $(\bar{y}_{Ney})$ **காணுதல் :**

$$\hat{V}\ (\bar{y}_{Ney}) = \frac{1}{N^2} \cdot \Sigma\, \frac{N_h - n_h}{N_h} \cdot \frac{(N_h\, s_h)^2}{n_h}.$$

$(N_h \cdot s_h)^2$	$\frac{(1)}{n_h}$	$\frac{(N_h - n_h)}{N_h}$	
(1)	(2)	(3)	(2) × (3)
$(960)^2$	$\frac{(960)^2}{29}$	$\frac{51}{80}$	20000
$(1280)^2$	$\frac{(1280)^2}{39}$	$\frac{121}{160}$	31000
$(1040)^2$	$\frac{(1040)^2}{32}$	$\frac{228}{260}$	30000
			81000

$$\therefore \quad \hat{V}\ (\bar{y}_{Ney}) = \frac{1}{(500)^2} \cdot (81000) = 0{\cdot}32$$

எடுத்துக்காட்டு 2 :

கொடுக்கப்பட்ட பட்டியலைப் பயன்படுத்தி, n-ன் மதிப்பை அறிந்து, ± 3 திட்ப அளவுப்படி, 95% நப்பகத் தன்மையுடன், அதனைப் பங்கீடு செய்க. வீதப் பங்கீட்டைப் பயன்படுத்துக.

படுகை	N_h	s_h	$\bar{x}_h$	$N_h s_h^2$
1	600	10	40	60000
2	300	20	60	120000
3	100	30	80	90000

95% நம்பகத் தன்மையில் $z = 2 \quad \therefore \quad z^2 = 4.$

$d = 3 \quad \therefore \quad d^2 = 9.$ எனவே $\dfrac{d^2}{z^2} = \dfrac{3}{2}$.

வீதப் பங்கீடு முறைப்படி,

$$n = \frac{N \cdot \Sigma N_h s_h^2}{\dfrac{N^2 \cdot d^2}{z^2} + \Sigma N_h s_h^2} = \frac{(1000)(270000)}{\dfrac{(100\ )^2 \cdot 9}{4} + 270000}$$

$\therefore \quad n = 108$ என்று மதிப்பிடப் படுகிறது.

எனவே,

$$n_h = \frac{N_h}{N} \cdot n$$

அதாவது,

$$n_1 = \frac{600}{1000} \times 108 = 65;$$

$n_2 = 32$; $n_3 = 11$ என்று பங்கீடு அமைகிறது.

$\therefore \quad n_1 + n_2 + n_3 = 108 = n$

பயிற்சி

நேமன் பங்கீட்டின்படி n எவ்வாறு பங்கு பிரிக்கப்படும் என்று கணக்கிடுக, அப்போது $\hat{V}(\bar{y}_{Ney})$ என்ன என்று தீர்மானிக்கவும்.

சென்ற பயிற்சியில், (4), (5), (6), (7) கணக்குகளில், நேமன் முறைப்படி n-ஐப் பங்கிட்டு, $\hat{V}(\bar{y}_{Ney})$ மதிப்பீட்டினைப் பெறுக.

படுகைப்படுத்திய ராண்டம் மாதிரி முறையில், திரும்பவைக்கும் வழிமுறை (Statisfic SRSWR) — With Replacement

h-ம் படுகையிலிருந்து, N_h கூறுகளைக் கருதி, n_h மாதிரிக் கூறுகள் மேற்கண்ட முறையில் தெரிவு செய்யப்படுகையில்.

$$\bar{y}_{st} = \Sigma w_h \cdot \bar{y}_h$$

என்று $\bar{Y}$-ன் பிறழ்ச்சியற்ற மதிப்பீடும், அதன் மாதிரி முறைசார் மாறுபாடானது.

$$V(\bar{y}_{st}) = \sum_h w_h^2 \cdot \frac{\sigma_h^2}{n_h}$$

எனவும் பெறப்படும். $\because \quad V(\bar{y}) = \frac{\sigma^2}{n}$ (SRSWR)

இதன் பிறழ்ச்சியற்ற மதிப்பீடு.

$$v(\bar{y}_{st}) = \Sigma w_h^2 \cdot \frac{s_h^2}{n_h} \text{ ஆகும்.}$$

$$\left[s_h^2 = \frac{1}{n_h - 1} \Sigma (y_{hi} - \bar{y}_h)^2 \right]$$

உத்தமப் பங்கீடு

$\bar{y}_h$-ன் மாதிரி முறை சார்ந்த மாறுபாடு, σ_h^2 ஆகும். எனவே, செலவு மாரு நிலையில் (c),

$$n_h = \frac{(c - c_0) \cdot \dfrac{N_h \sigma_h}{\sqrt{c_h}}}{\sum_h N_h \sigma_h \cdot \sqrt{c_h}}$$

$$= \frac{(c - c_0) \left(\dfrac{w_h \sigma_h}{\sqrt{c_h}} \right)}{\Sigma w_h \sigma_h \cdot \sqrt{c_h}}$$

V என்பது மாருத அளவாக இருப்பின், உத்தம n_h மதிப்பினை $V(\bar{y}_{st})$ யில் பிரதியிட்டு,

$$V = V(\bar{y}_{st}) = \Sigma \frac{w_h^2 \cdot \sigma^2_h}{n \cdot \frac{w_h \cdot \sigma_h}{\sqrt{c_h}}} \cdot \Sigma \frac{w_h \cdot \sigma_h}{\sqrt{c_h}}$$

$$= \frac{1}{n} \left(\sum_h w_h \, \sigma_h \cdot \sqrt{c_h} \right) \left(\sum_h \frac{w_h \, \sigma_h}{\sqrt{c_h}} \right)$$

அல்லது

$$n = \frac{1}{V} \left(\sum_h w_h \, \sigma_h \sqrt{c_h} \right) \left(\sum_h \frac{w_h \, \sigma_h}{\sqrt{c_h}} \right)$$

எல்லாப் படுகை கட்கும், $c_h = c'$ ஆனால்,

$$n_h = n \cdot \frac{w_h \cdot \sigma_h}{\sum_h w_h \cdot \sigma_h}$$ என்றாகும்.

இதில்

$$n = \frac{c - c_0}{c'} .$$

இதற்கேற்ற, மாதிரி முறை சார்ந்த மாறுபாடானது.

$$V_{min}(\bar{y}_{st}) = \frac{1}{n} \left(\sum_h w_h \, \sigma_h \right)^2 \qquad \text{......(1)}$$

(சிறும அளவில் அமைந்த மாறுபாடு V_{min} ஆகும்).

முழுமைத் தொகுதியிலிருந்து, திரும்பவைக்கும் எளிய சம வாய்ப்பு முறைப்படி (SRSWR), படுகைப்படுத்தாத நிலையில், மாதிரிச் சராசரியின் மாறுபாடானது

$$V(\bar{y}_{us}) = \frac{\sigma^2}{n}$$

(us என்பது படுகையாக்கப்படாத நிலையினைக் காட்டும்)

இதில், $\sigma^2 = \frac{1}{N} \Sigma \, \Sigma \, (y_{hi} - \bar{y})^2$

மேலும்,

$$V(\bar{y}_{us}) = \frac{1}{N} \sum_h W_h \, \sigma_h^2 + \frac{1}{n} \Sigma \, w_h . \, (\bar{y}_h - \bar{Y})^2 \qquad \text{...... (2)}$$

(ஏற்கெனவே அறிந்தபடி).

(1), (2) இவற்றை ஒப்பிட்டு,

$$V_{min}(\bar{y}_{st}) = V(\bar{y}_{us}) - \frac{1}{n}\sum_h w_h.(\bar{Y}_h - \bar{Y})^2$$

$$-\frac{1}{n}\sum_h w_h\ (\sigma_h - \bar{\sigma})^2$$

என்று பெறலாம். $\bar{\sigma} = \Sigma\, w_h.\sigma_h$.

எனவே படுகையாக்கப்படாத எளிய சமவாய்ப்பு முறையினை விட அதிக நன்மை (gain) அல்லது இலாபம் பெற, படுகை சராசரிகளிடையேயும், மற்றும் படுகைத் தரவிலக்கங்களுக்கிடையேயும் அதிக அளவில் மாறுபாடு அல்லது வேறுபாடு ஏற்படும் அளவில், படுகைபிரித்தல் அமைதல் வேண்டும்.

விதப் பங்கீடு :

$n_h = n.w_h$ என்ற அடிப்படையில், வீதப் பங்கீட்டு முறைப்படி $V(\bar{y}_{st})$ யானது,

$$V_{prop}(\bar{y}_{st}) = \frac{1}{n}\Sigma\, w_h \sigma_h^2$$

V_{prop} என்பது வீதப்பங்கீட்டைக் குறிக்கும்.

(1) மற்றும் (2) என்று மேலே குறிக்கப்பட்ட மாறுபாட்டு அளவுகளுடன் இதை ஒப்பிட,

$$V_p(\bar{y}_{st}) = V_{min}(\bar{y}_{st}) + \frac{1}{n}\sum_h w_h\ (\sigma_h - \bar{\sigma})^2$$

என்றும்,

$$V_p(\bar{y}_{st}) = V(\bar{y}_{us}) - \frac{1}{n}\sum_h w_h\ (\bar{Y}_h - \bar{Y})^2 \qquad \ldots\ldots \quad (4)$$

என்றும் பெறலாம்.

(4)-ன் மூலமாக, படுகைச் சராசரிகள் ஒன்றுடன் ஒன்று அதிக அளவில் மாறுபடும் வகையில் படுகைப் படுத்துகையில், வீதப் பங்கீட்டு முறை அதிகப் பயன் அளிக்கிறது எனக் காணலாம்.

படுகையாக்குவதில் உள்ள பலன் அல்லது ஆதாயம் (Gain due to Stratification)

$$V(\bar{y}_{st}) = \sum_h w_h^2 . \frac{\sigma_h^2}{n_h}$$

என்பதையும்,

(2) என்று குறிக்கப்பட்ட $V(\bar{y}_{us})$ மதிப்பினையும் ஒப்பிடுகையில் படுகையாக்குவதன்மூலம் அல்லது பகுதிப் பிரிப்பின் மூலம் ஏற்படக்கூடிய மாறுபாட்டுக் குறைவு (reduction in sampling variance) ஆனது,

$$V(\bar{y}_{us}) - V(\bar{y}_{st}) = \frac{1}{n} \sum_h \left(1 - \frac{n.w_h}{n_h}\right). \; w_h . \sigma_h^2$$

$$+ \frac{1}{n} \sum_h w_h . (\bar{Y}_h - \bar{Y})^2 \quad \ldots\ldots \quad (5)$$

என்றிருக்கக் காணலாம். எனவே, படுகைமுறை மாதிரித் தெரிவானது, படுகை பிரிக்கப்படாத எளிய ராண்டம் மாதிரித் தெரிவினைவிட, அதிகத் திறனுடையதாக அறியலாம். இதற்கேற்ப, n_h-அளவானது (i) w_hக்கு வீதத் தொடர்பிலோ அல்லது (ii) $w_h\sigma_h$ க்கு வீதத் தொடர்பிலோ (proportional) எடுக்கப்படல் வேண்டும். வலப்புறம் முதல் உறுப்பானது, n_h-அளவு $w_h\sigma_h^2$ உடன் வீதத் தொடர்பில் எடுக்கப்பெறுகையில், பூச்சியமாக ஆனாலும், அதே அளவு மாதிரிமுறை சார்ந்த மாறுபாட்டு அளவினை, வீதப் பங்கீட்டு முறைப்படி பெறமுடிகிறதென்பது குறிப்பிடத்தக்கது. அது எளிதும் ஆகும்.

குறிப்பிட்ட மூன்று முறைகளில் அல்லாது வேறு வழியில் பங்கீடு செய்யப்படுமானால், இந்த உறுப்பானது எதிர்க்கணியமாக (— ve) அல்லது மற்ற உறுப்பினைக் காட்டிலும் அதிக அளவினதாய் அமைய, திறன் இழப்பு ஏற்படும்.

∴ படுகைகளுக்குப் பொருந்திய அளவில் பங்கீடு அமைதல் இன்றியமையாதது.

படுகை முறையில் ஏற்படும் மாறுபாட்டு அளவின் குறைவினை (reduction), பிறழ்ச்சியற்று மதிப்பீடு செய்தற்கு, முதன்முதலில்

$V(\bar{y}_{us})$-ன் பிறழ்ச்சியற்ற மதிப்பீட்டை பெறவேண்டும். இது, படுகை பிரிக்கப்படாத அல்லது பகுக்கப்படாத மாதிரியின் அடிப்படையில் அமையும். அதாவது, σ^2 மதிப்பை பிறழ்ச்சியற்று மதிப்பிட்டு அறியலாம்.

$$\sigma^2 = \frac{1}{N}\sum_h \sum_i (y_{hi}^2) - (\bar{Y})^2.$$

முதல் உறுப்பின் பிறழ்ச்சியற்ற மதிப்பீடாவது :

$$\sum_h \frac{w_h}{n_h} \sum_i y_{hi}^2$$

$\bar{Y}^2$ ஆனது $\bar{y}_{st}^2 - v(\bar{y}_{st})$ மூலமாக பிறழ்ச்சியற்று மதிப்பிடப் படுகிறது. $v(\bar{y}_{st})$ யானது, $V(\bar{y}_{st})$ யின் பிறழ்ச்சியற்ற மதிப்பீடாகும்.

$$\therefore \quad v(\bar{y}_{us}) = \frac{1}{n}\left[\sum_h \frac{w_h}{n_h} \sum_i y_{hi}^2 - \bar{y}_{st}^2 + v(\bar{y}_{st})\right]$$

ஆகும்.

படுகை மாதிரிமுறை மற்றும் எளிய ராண்டம் மாதிரிமுறை—ஓர் ஒப்புமை (Stratified and Simple Random Sampling)

தகுந்த முறையில் படுகைகள் பிரிக்கப்பட்டு, பங்கீடும் சரிவர அமைந்திடின், படுகைமாதிரிமுறை அதிக திட்பம் வாய்ந்ததாக அமைந்திடும். மதிப்பீட்டுத் திறனை (efficiency) ஒப்பிடுகையில், இருமுறைகளில் பெறப்படும் மதிப்பீட்டு அளவைகளின் மாறுபாட்டுத் தன்மையை (variances of estimators) ஒப்பிடுதல் பொருந்தும்.

தேற்றம் 7 :

$\frac{n_h}{N_h}$ என்ற விகிதம் வரும் உறுப்புக்களை நீக்கிவிட்டு,

$$V_{opt} < V_{prop} \leq V_{ran}$$

என்ற உண்மையினைப் பெறலாம் உத்தமப் பங்கீடானது, n-மதிப்பை மாறுததாக வைத்துக்கொண்டு கருதப்படுகிறது. மதிப்பீட்டுச் சராசரிகளின் மாறுபாட்டினை, ஒவ்வொரு முறைக் கேற்ப, V_{opt}, V_{prop}, V_{ran} என்று குறிப்பிடுகிறோம். V_{ran} என்பது சமவாய்ப்பு மாதிரிமுறையில் உள்ள மதிப்பு.

நிரூபணம் :

fpc காரணியினைக் கருதாது,

$$V_{ran} = \frac{S^2}{n}$$

$$V_{prop} = \frac{\Sigma N_h S_h^2}{n N}$$

$$V_{opt} = \frac{(\Sigma N_h S_h)^2}{n.N^2}$$

என எழுதலாம்.

படுகை முறைப்படி,

$$(N-1) S^2 = \sum_h \sum_i (y_{hi} - \bar{Y})^2$$

$$= \sum_h \sum_i (y_{hi} - \bar{Y}_h)^2 + \sum_h N_h (\bar{Y}_h - \bar{Y})^2$$

$$= \sum_h (N_h - 1) S_h^2 + \sum_h N_h (\bar{Y}_h - \bar{Y})^2$$

$\left\{ \frac{1}{N_h} \right.$ உள்ள உறுப்புகள் மிகச் சிறியனவாதலால், இவ்வாறு எழுதுவது சாத்தியமாகிறது $\}$.

எனவே $$NS^2 = \sum_h N_h S_h^2 + \sum_h N_h (\bar{Y}_h - \bar{Y})^2$$

இதிலிருந்து $$V_{ran} = \frac{S^2}{n} = \frac{\Sigma N_h S_h^2}{n.N} + \frac{\Sigma N_h (\bar{Y}_h - \bar{Y})^2}{n.N}$$

அதாவது $$V_{ran} = V_{prop} + \frac{\Sigma N_h (\bar{Y}_h - \bar{Y})^2}{nN}$$

இதன்மூலம்

$$V_{prop} \leq V_{ran}$$

எனப் பெறப்பட்டு,

வீத அடிப்படையில் அமையும் படுகை மாதிரிமுறை மாறுபாடு, ராண்டம் மாதிரிமுறை மாறுபாட்டைவிடக் குறைவு என முடிவு செய்யப்படும்.

அடுத்து, வரையறையின்படி,
$V_{prop} \geq V_{opt}$ என்றிருத்தல் வேண்டும்.

$fpc = 1$ எனக் கொண்டு,

$$V_{prop} = \frac{\Sigma N_h S_h^2}{Nn}$$

$$= \sum_h \frac{N_h S_h^2}{Nn} + \frac{\left[\sum_h N_h S_h\right]^2}{n \cdot N^2} - \frac{\left[\sum_h (N_h S_h)\right]^2}{nN^2}$$

$$= V_{opt} + \frac{1}{Nn}\left[\sum_h N_h s_h^2 - \frac{\left(\sum_h N_h s_h\right)^2}{N}\right]$$

$$\therefore \quad V_{prop} - V_{opt} = \frac{1}{Nn} \Sigma N_h (S_h - \bar{S})^2$$

என்றாகும். இதில் $\bar{S} = \sum_h \frac{N_h S_h}{N}$

$$\left\{ \frac{1}{nN} \Sigma N_h (S_h - \bar{S})^2 = \frac{1}{nN}\left[\Sigma N_h S_h^2 - \frac{(\Sigma N_h S_h)^2}{N}\right] \right.$$

என்று காட்டுக. $\}$

எனவே,

$$V_{ran} = V_{opt} + \frac{\sum_h N_h (S_h - \bar{S})^2}{nN} + \frac{\Sigma N_h (\bar{Y}_h - \bar{Y})^2}{nN}$$

என்பதன் மூலமாக,

$$V_{opt} \leq V_{st} \leq V_{ran}$$

என்பது புலனாகிறது. இது உத்தமப் பங்கீட்டின் திட்ப அளவினை நன்கு விளக்குகிறது.

வீத அளவு பற்றிய படுகை மாதிரிமுறை (ராண்டம்முறை) (Stratified Random Sampling for Proportions)

முழுமைத் தொகுதியின் வீத அளவு (proportion) பற்றிக் கருதுகையில், படுகைமுறையானது திறனுடன் பயன்படுத்த படலாம். உதாரணமாக, ஒரு ஊரின் மக்கட் தொகையில் வயதின் அடிப்படையில் படுகைகளை அமைத்து, அதன்மூலமாக உடல் ஊனமுற்றோர் (ஏதேனும் ஓர் குறிப்பிட்டவகையில்) வீத அளவினை மதிப்பீடு செய்வதாகக் கொள்ளலாம். இத்தகைய நிலைகளில், வீத மதிப்பீடுகளைப் பெறுவதே கருதப்படும் பிரச்சினையாகும்.

முழுமைத் தொகுதி வீதமானது (proportion) P என்று கருதப் பட்டு, துணைத் தொகுதி வீதங்கள் P_1, P_2... என்று குறிக்கப்பட்டு, படுகைமுறையில் சமவாய்ப்பு மாதிரித் தேர்வு அடிப்படையில், y என்ற மாறியினைக் கருதி, முன்பு போலவே முடிவுகளைப் பெறலாம்.

P-யின் மதிப்பானது, $\bar{Y}$ மதிப்பினைப்போலவே பெறப்படும். P_h என்பதை h-ம் படுகையின் வீதமாகக் கருதி,

$$P_h = \bar{Y}_h = \frac{1}{N_h} \sum_{1}^{N_h} y_{hi}$$

என்று பெறப்படும்.

($y_{hi} = 0$ அல்லது 1 என்று கொள்ளப்படும்)

$$\therefore \quad P = \frac{\sum_{h} N_h \, P_h}{N}$$

$$= \frac{\sum_{h} \sum_{i} y_{hi}}{N} \quad \text{ஆகும்.}$$

மேலும் $\hat{Y} = \bar{y}_{st} = \frac{\Sigma N_h \bar{y}_h}{N}$ என்பது $\bar{Y}$-ன் பிறழ்ச்சியற்ற மதிப்பீடாக, அதன் அடிப்படையில்,

$$\hat{P} = p_{st} = \frac{\sum_h N_h \cdot p_h}{N}$$

என்று பெறப்படும்.

இதில் $p_{st} = \frac{1}{n_h} \sum_i y_{hi}$

மேலும்

$$E(p_h) = \frac{1}{n_h} \sum_i^{n_h} E(y_{hi})$$

$$= \frac{1}{n_h} \cdot n \cdot \sum_i^{N_h} \frac{1}{N_h} \cdot y_{hi}$$

$$= \frac{1}{N_h} \sum_i y_{hi} = P_h$$

$$\therefore \quad E(p_{st}) = \frac{\Sigma' N_h \cdot E(p_h)}{N} = \frac{\Sigma N_h P_h}{N} = P \text{ ஆகும்.}$$

$\therefore$ p_{st} ஆனது P-யின் பிறழ்ச்சியற்ற மதிப்பீடாக அறியப் படுகிறது.

p_{st} மதிப்பீட்டின் மாதிரி முறைசார் மாறுபாடு (Sampling Variance of p_{st})

$y_i = 0$ அல்லது 1 எனக் கொண்டு, N அளவில் உள்ள முழுமைத் தொகுதியின் மாறுபாடு

$$S^2 = \frac{N}{N-1} \cdot PQ \text{ ஆகும்.}$$

$Q = (1 - P)$.

படுகை மாறுபாடுகள் S_h^2 எனப்பட்டு,

$$S_h^2 = \frac{N_h}{N_h - 1} \cdot P_h \, Q_h$$

என்று அமையும். $(Q_h = 1 - P_h)$.

முன்பு போலவே, $V(\bar{y}_{st})$ என்பது, இந் நிலையில் $V(p_{st})$ என்று ஆக, $\{ y_i = 0$ அல்லது,$\}$

$$V_{(pst)} = \frac{1}{N^2} \sum_h N_h^2 \cdot \frac{N_h - n_h}{N_h} \cdot \frac{1}{n_h} \cdot \frac{N_h}{N_h - 1} \cdot P_h \cdot Q_h$$

என்று நிறுவப்படும்.

fpc-ஐப்புறக்கணிக்கத்தக்க அளவில்,

$$V_{(pst)} = \frac{1}{N^2} \Sigma N_h^2 \cdot \frac{P_h Q_h}{n_h}$$

என்று மதிப்பிடப்படும்.

இதன் மாதிரி மதிப்பீடு,

$$\hat{V}_{(pst)} = \frac{1}{N^2} \sum_h N_h^2 \cdot \frac{N_h - n_h}{N_h} \cdot \frac{1}{n_h} \cdot \frac{N_h}{N_h - 1} \, p_h \cdot q_h$$

இது $V_{(pst)}$-யின் பிறழ்ச்சியற்ற மதிப்பீடாகும்.

வீதப் பங்கீடு

$n_h = \frac{N_h}{N} \cdot n$ எனும் அடிப்படை.

உத்தமப் பங்கீடு

மொத்த செலவு நிலையானதாக, மாறுபாடு குறைக்கப்படும் அளவில், அடிப்படையாவது :

$$n_h = \frac{N_h \, S_h / \sqrt{c_h}}{\Sigma N_h \, S_h / \sqrt{c_h}} \cdot n$$

இதில்

$S_h^2 = \frac{N_h}{N_h - 1} \cdot P_h \, Q_h \doteqdot P_h \, Q_h$ என்று பிரதியிட,

$$n_h = \frac{N_h \cdot \frac{\sqrt{P_h Q_h}}{\sqrt{c_h}}}{\Sigma \left(N_h \cdot \frac{\sqrt{P_h Q_h}}{\sqrt{c_h}} \right)} \cdot n$$

என்று அமைகிறது. p_h, q_h என்ற மதிப்பீடுகள் நடைமுறையில் பயன்படுத்தப்படுகின்றன.

நேமன் பங்கீடு

c_h மாறிலியாக, எனவே n-ம் மாறிலியாக,

$$n_h = \frac{N_h \cdot \sqrt{P_h Q_h}}{\Sigma N_h \cdot \sqrt{P_h Q_h}} \cdot n \text{ என்று ஆகும்.}$$

எடுத்துக்காட்டு 1 :

	y மதிப்புகள் :				
படுகை 1 :	1	0	1	0	0
படுகை 2 :	1	0	1	1	1

இதில்,

$$P_1 = \frac{\left[\sum_i y_{1i}\right]}{N_1} = \frac{2}{5}$$

$$P_2 = \frac{\left[\sum_i y_{2i}\right]}{N_2} = \frac{4}{5}.$$

$$N_1 = 5 ; \quad N_2 = 5.$$

$$\therefore \quad P = \frac{N_1 P_1 + N_2 P_2}{N_1 + N_2} = \frac{5 \cdot \left(\frac{2}{5}\right) + 5 \cdot \left(\frac{4}{5}\right)}{10}$$

$$= \frac{6}{10} = 0 \cdot 6.$$

இதிலிருந்து எடுக்கப்படும் (1, 0), (1, 1, 0) என்ற மாதிரிகளைக் கருதுவோம்.

அவற்றில், $p_1 = \frac{1}{2}$; $p_2 = \frac{2}{3}$.

$$\therefore \quad \hat{P}_1 = p_1 = 0{\cdot}5 ; \quad \hat{P}_2 = p_2 = 0{\cdot}67$$

P-யின் மதிப்பீட்டு அளவை :

$$p_{st} = \Sigma \frac{N_h \, p_h}{N} = \frac{(5 \times 0{\cdot}5) + (5 \times 0{\cdot}67)}{10}$$

$$\therefore \quad p_{st} = 0{\cdot}585$$

பயிற்சி :

படுகை 1 :	1	0	1
படுகை 2 :	1	1	0

இவற்றிலிருந்து $n_1 = 2$; $n_2 = 2$ அளவில் மாதிரிகள் எல்லா வற்றையும் கருதி, $E(p_{st}) = P$ என நிறுவுக.

எடுத்துக்காட்டு 2 :

நான்கு பிரிவுகளில் அமைந்த மாணவர்கள் விவரம் வருமாறு:

பிரிவு	N_h	n_h	p_h	$N_h \cdot p_h$
1	200	50	0·2	40
2	160	40	0·3	48
3	120	30	0·4	48
4	120	30	0·5	60
	600			196

p_h அளவுகள் ஏற்கெனவே தீர்மானிக்கப் பெற்றவை.

$$\therefore \quad p_{st} = \frac{1}{N} \Sigma N_h p_h$$

$$= \frac{1}{600}(196) = 0{\cdot}33$$

மேலும்

$$S_h^2 = \frac{N_h}{N_h - 1} \cdot P_h \cdot Q_h$$

$$V(p_{st}) = \frac{1}{N^2} \Sigma N_h^2 \cdot \frac{N_h - n_h}{N_h} \cdot \frac{1}{n_h} \cdot \frac{N_h}{N_h - 1} \cdot P_h \cdot Q_h$$

என்பது p_{st}-யின் மாதிரி முறை சார் மாறுபாடு.

ஆனால் இங்கு கொடுக்கப்பட்டவை P_h-ன் மதிப்பீடுகள்.

எனவே,

$$\hat{V}(p_{st}) = \frac{1}{N^2} \Sigma N_h (N_h - n_h) \cdot \frac{1}{n_h} \cdot p_h \cdot q_h$$

$$(q_h = 1 - p_h).$$

N_h	n_h	p_h	$p_h q_h$	$(N_h - n_h)$	$\left[N_h \cdot \frac{(N_h - n_h)}{n_h} \cdot p_h q_h\right]$
200	50	0·2	0·16	150	96·0
160	40	0·3	0·21	120	100·8
120	30	0·4	0·24	90	86·4
120	30	0·5	0·25	90	90·0
					373·2

$$\therefore \hat{V}(p_{st}) = \frac{1}{(600)^2}[373{\cdot}2]$$
$$= 0{\cdot}0010 \ldots\ldots$$
$$\therefore S(p_{st}) = \sqrt{\cdot 001} = 0{\cdot}03$$

திரும்பவைக்கும் முறை :

(*i*) வீதப் பங்கீட்டு மாறுபாடு

$$V(p_{st}) = \frac{1}{n}\Sigma w_h P_h Q_h$$

(*ii*) உத்தமப் பங்கீட்டு மாறுபாடு

$$V_{opt} = \frac{1}{n}\left(\Sigma w_h \sqrt{P_h Q_h \cdot c_h}\right)\left(\Sigma w_h \frac{\sqrt{P_h Q_h}}{c_h}\right)$$

(*iii*) நேமன் பங்கீடு :

$$V_{Ney} = \frac{1}{n}(\Sigma w_h \sqrt{P_h Q_h})^2.$$

தொடர்ச்சியான விவரத்தின்மூலம், மாதிரி அளவு மதிப்பீடு (Estimation of Sample Size with Continuous Data)

மதிப்பீட்டு அளவானது (estimate) ஒரு குறிப்பிட்ட மாறுபாடு V-யினை உடையதாகக் கருதப்படுகிறது. பிழை எல்லை (margin of error) d என்றமையும் நிலையில் $V = \left(\frac{d}{z}\right)^2$ என்றாகும். இதில் z என்பது, பிழையானது குறிப்பிட்ட எல்லையினைக் கடக்க அனுமதிக்கப்படும் வாய்ப்புக்கேற்ப உள்ள இயல்நிலை விலக்க அளவு.

$\bar{Y}$-ன் மதிப்பீடு:

s_h என்பது S_h-ன் மதிப்பீடாகவும், $n_h = w_h \cdot n$ என்றும் கொள்க. (w_h தெரிவு செய்யப்பட்டவை) இந் நிலையில், எதிர் பார்க்கும் மாறுபாடு,

$$V(\bar{y}_{st}) = \frac{1}{n} \Sigma \frac{W_h^2 . s_h^2}{w_h} - \frac{1}{N} \Sigma W_h \cdot s_h^2$$

$$\left(W_h = \frac{N_h}{N} \right) \cdot$$

இதன்மூலம்

$$n = \frac{\sum\limits_h \dfrac{W_h^2 \cdot s_h^2}{w_h}}{V + \dfrac{1}{N} \Sigma W_h s_h^2}$$

என்று பெறலாம்.

fpc-யினை நீக்கிவிட்டு, முதல் தோராய மதிப்பாக,

$$n_0 = \frac{1}{V} \Sigma \frac{W_h^2 \cdot s_h^2}{w_h}$$ என்று கொள்வோம்.

$\frac{n_0}{N}$ என்பது புறக்கணிக்கப்படாத நிலையில்,

$$n = \frac{n_0}{1 + \dfrac{1}{N \cdot V} \Sigma W_h s_h^2}$$

என்று கணக்கிடப் பெறும்.

குறிப்பு : நடைமுறையில், *fpc* தக்கத்தள்ள அளவினதாய் அமையாவிடினும், $\frac{1}{N_h}$ உள்ள உறுப்புகள் மிகச் சிறிய அளவினதாக அமையும். எனவே,

திரும்ப வைக்கும் முறையில், படுகை சமவாய்ப்பு மாதிரித் தேர்வில்,

$$(Vp_{st}) = \frac{1}{N^2} \Sigma N_h (N_h - n_h) \cdot \frac{P_h Q_h}{n_h}$$

$$\therefore \quad (Vp_{st}) = \sum_h w_h^2 \cdot \frac{P_h Q_h}{n_h} (1 - f_h)$$

என்றும், fpc புறக்கணிக்கப்படின்

$$V(p_{st}) = \Sigma w_h^2 \cdot \frac{P_h Q_h}{n_h}$$

என்றும் பயன்படுத்துகிறோம்.

திரும்ப வைக்காத முறையில் (Without Replacement),

(*i*) **வீதப் பங்கீடு**

$$V_{prop}(p_{st}) = \frac{N-n}{N} \cdot \frac{1}{nN} \cdot \Sigma N_h^2 \cdot \frac{P_h Q_h}{N_h - 1}$$

$$= \frac{1-f}{n} \Sigma W_h \cdot P_h Q_h \cdot$$

(*ii*) **நேமன் பங்கீட்டில்**

n மாறுததாக,

$$n_h = \frac{w_h \cdot \sqrt{P_h Q_h}}{\Sigma w_h \sqrt{P_h Q_h}} \cdot n$$

என்றும்,

$$V(p_{st}) = \frac{1}{n}(\Sigma w_h \sqrt{P_h Q_h})^2 - \frac{1}{N} \Sigma w_h P_h Q_h$$

(*iii*) **உத்தம பங்கீட்டில்,**

$$n_h = \frac{N_h \sqrt{P_h Q_h / c_h}}{\Sigma N_h \sqrt{P_h Q_h / c_h}}$$

என்றும்,

$$V_{opt}(p_{st}) = \frac{1}{n}(\Sigma w_h \sqrt{P_h Q_h c_h}) \Sigma w_h \sqrt{P_h Q_h / c_h} - \sum_h \frac{w_h^2 \cdot P_h Q_h}{N_h}$$

என்றும் பெறுகிறோம்.

நேமன் பங்கீட்டில், $(w_h \propto W_h s_h)$.

$$n = \frac{(\Sigma W_h \cdot s_h)^2}{V + \frac{1}{N} \Sigma w_h s_h^2}$$

$$\left[V = D^2 = \frac{d^2}{z^2}\right]$$

வீதப் பங்கீட்டில்,

$$n_0 = \frac{\Sigma\, w_h\, s_h^2}{V} \text{ என்று கொண்டு,}$$

$$n = \frac{n_0}{1 + \frac{n_0}{N}} \text{ என மதிப்பிடப்பெறும்.}$$

பயிற்சி :

இவற்றை $\hat{Y}_{st}$-யின் மதிப்பீட்டு அடிப்படையில் கருதி, n-மதிப்புகளை வெவ்வேறு பங்கீடுகளில் பெறுக.)

வீத அளவுகளுடன், மாதிரி அளவு நிர்ணயித்தல் (Proportion & & Sample Size).

இங்கு மாறுபாடு V ஆனது, வீத அளவு (Proportion) P-யின் மதிப்பீடாக அமையும். இதைக் கருதிப் பின்வரும் மதிப்புகளைப் பெறலாம் :

வீதப்பங்கீட்டில்,

$$n_0 = \frac{\Sigma\, w_h\, p_h\, q_h}{V} \text{ என்று ஆவதால்,}$$

$$n = \frac{n_0}{1 + \frac{n_0}{N}} \text{ எனப் பெறப்படும்.}$$

நேமன் பங்கீட்டில்,

$$n_0 = \frac{\left(\sum_h w_h \sqrt{p_h\, q_h}\right)^2}{V};$$

$$\therefore \quad n = \frac{n_0}{1 + \frac{1}{NV} \cdot \sum_h w_h\, p_h\, q_h}$$

n_0 என்பது முதல் தோராய மதிப்பு (approximate).

$\left(\frac{N_h}{N_h - 1}\right.$ என்பது ஒன்றாகக் கருதப்படும்.$\left.\right)$

இதேபோல, ஒரு பிரிவில் உள்ள மொத்த முழுமைத் தொகுதிக் கூறுகளைக் கருதுகையில், (அதாவது N_p-யினை மதிப்பிடுகையில்), மாறுபாடுகள் N^2 காரணியினால் பெருக்கப்படவேண்டும்.

எடுத்துக்காட்டு

மூன்று நகர்களில் நடத்தப்பட்ட ஆய்வில், 2 அல்லது அதற்குமேல் குழந்தைகள் உள்ள குடும்பங்கள் கருதப்பட்டன.

விவரங்கள்:

நகர்	குடும்பங்கள்	p_h
1	2000	0·20
2	3000	0·30
3	5000	0·40

(i) P-யிலிருந்து $\pm 3\%$-ல் அமையும் திட்பத்துடன் n-ன் மதிப்பைப் பெறுக.

(ii) பெற்ற n-மதிப்பை, வீதப் பங்கீட்டில் பங்கிடுக.

(i) n-அளவு

$N_h \cdot p_h$	$N_h \cdot p. q_h$
400	320
900	270
500	200
1800	790

$$d = |p_{st} - P| \leq 0·03.$$

$z = 3$ எனக் கொள்க. (99% நம்பகத் தன்மை).

$$\therefore \quad D = \frac{d}{z} = \frac{0·03}{3} = 0·01 \,; \quad D^2 = (0·01)^2$$

வரையறைப்படி:

$$n = \frac{N \sum_h N_h S_h^2}{N^2 D^2 + \Sigma N_h s_h^2} = \frac{N \Sigma N_h p_h q_h}{N^2 D^2 + \Sigma N_h p_h q_h}$$

$(fpc \neq 1)$.

$$= \frac{(10000)(790)}{(1000 \times 0{\cdot}01)^2 + 790} \doteq 888.$$

$\therefore \; n = 888$ (தோராயமாக)

(*ii*) **வீதப் பங்கீடு**

$$n_1 = \frac{2000}{10000} \times 888 \doteq 178$$

$$n_2 = \frac{3000}{10000} \times 888 \doteq 266$$

$$\& \; n_3 = \frac{5000}{10000} \times 888 = 444$$

6. ஒழுங்கு மாதிரி முறை
(Systematic Sampling Method)

[அறிமுகம் — மாதிரித் தெரிவுமுறை — சராசரி மதிப்பீடும், அதன் மாறுபாட்டு அளவும் — எடுத்துக்காட்டுகள்— ஒழுங்கு முறையில் சில நிலைகள் — பயிற்சி.]

அறிமுகம்

மாதிரியினைத் தெரிவு செய்யும்பொருட்டுக் கருதப்படும் முழுமைத் தொகுதியின் கூறுகள் (units) பற்றிய விவரப் பட்டியல் கொடுக்கப்பட்ட நிலையில், ஒரு முறையான அடிப்படையில் மாதிரித் தேர்வு நடத்தப்படுவது, செயல்முறைக்கு எளிதாக அமைவதைக் காணலாம். பட்டியலில், சமமான இடைவெளியில் அமையப்பெற்ற கூறுகள் ஒரு தெளிவான முறையின் அடிப் படையில் எடுக்கப்பெறுகின்றன. உதாரணமாக, N அளவுள்ள தொகுதியை 1 முதல் N வரை இலக்கமிட்டு, அதிலிருந்து n-அளவுள்ள மாதிரிக்கூறு எடுக்கவேண்டியிருப்பின், $N \leq nK$ என்ற விதிப்படி ஒவ்வொன்றிலும் K-கூறுகள் அமையும்படியாக முழுமைத் தொகுதியானது பிரிக்கப்பட்டு, ஒரு குறிப்பிட்ட உறுப்பினில் துவங்கி, அதிலிருந்து ஒவ்வொரு K-ஆம் உறுப்பும் தெரிவு செய்யப்பட்டு மாதிரிக்கூறு அமைக்கப்படுகிறது.

I என்ற முதல் உறுப்பு, வாய்ப்பு அடிப்படையில் தெரிவு செய்யப்படின், (1-லிருந்து K வரை உள்ள எண்களில் ஏதேனும் ஒன்று), அதன்பின் $(I+K)$, $(I+2K)$......... $[I+(n-1.)K]$ என்ற உறுப்புகள் தெரிவு செய்யப்பட்டு, மாதிரியானது பெறப் படும். பல்வேறு நிலைகளில் இம்மாதிரிமுறை சிறந்த பலனை அளிக்கிறது. உதாரணமாக, மக்கட்தொகைக் கணிப்புபற்றிய ஆய்வுகளில், காலவிரயத்தைத் தவிர்க்க, சில உடனடி முடிவுகள் பெறும் பொருட்டு இம் முறை மேற்கொள்ளப்படலாம்.

மாதிரி எடுத்தல் எளிதாக இருப்பதால், உடனடியாகச் சில முடிவுகளைப் பெறுவது இயலுகிறது. நடைமுறையில், ஏதேனும் ஒரு குறிப்பிட்டப் பிரச்னையைப் பற்றிப் பொதுமக்கள் கருத்தினை அறிவதற்கு மேற்கொள்ளப்படும் ஆய்வுகளில், முறையுடை (systematic) அடிப்படையானது சிறந்த பலன் தருகிறது.

எளிதில் புரிந்துகொள்ளக்கூடிய ஒரு வழிமுறையாகவும் அமைகிறது. கருதப்படும் பகுதியில் (area), எடுக்கப்படும் நபர்களை வரிசைப்படுத்தி இலக்கமிட்டு, அவற்றிலிருந்து சில கூறுகளை (units) ஒரு குறிப்பிட்ட நியதி அல்லது ஒழுங்கு அடிப் படையில் மாதிரித் தெரிவு நடைபெறுகிறது. வானொலி கேட்போ-ரிடையே, ஒரு குறிப்பிட்ட நிகழ்ச்சியினைப் பற்றிக் கருத்தினை ஆராயும்பொருட்டு, வானொலி நிலையத்தினர் மேற்கொள்ளும் ஆய்வுகளில் இம்முறையானது அதிக அளவில் பயன்படக் காணலாம். மாதிரிக் கூறுகள் (sampling units), முழுமைத் தொகுதியினிடையே பரவலாகக் காணப்படும் நிலையில், இம்முறை யானது சிறந்த மாதிரியினைத் தர இயலுகிறது. திட்பமும் அதிக மாவதைப் பின்பு காண்போம். சில நேரங்களில், சமவாய்ப்பு முறையினைவிட, இம்முறைத் தேர்வானது விரும்பப்படலாம்.

பொதுவாக, மக்கட்தொகை கணிப்பு பற்றிய ஆய்வுகளில், உடனடியாக சில உண்மைகளை அறிவதற்காக, சிறு அளவில் துணைக் கூறுகள் (sub-samples) கருதப்பட்டு அவற்றின் அடிப் படையில் மதிப்பீடுகள் பெறப்படுகின்றன. இம் மதிப்பீடுகள் திறனுடனும் அமைந்திருக்கும்.

ஒரு தொகுதியிலிருந்து, ஏதேனும் ராண்டமாகத் தெரிந் தெடுக்கப்பட்ட உறுப்பிலிருந்து தொடங்கி, அதிலிருந்து ஒவ்வொரு k-ஆம் உறுப்பினையும், ஓர் ஒழுங்கின் அடிப்படையில் தெரிவு செய்வது இம்முறை. உதாரணமாக, 1, 2 ....N என்று இலக்கமிட்ட தொகுதியில், 5வது உறுப்பில் இருந்து துவங்கி, ஒவ்வொரு 10ஆவது உறுப்பினையும் தெரிந்தெடுப்பதாகக் கொண்டால், 5, 15, 25 ..... என்ற வகையில் மாதிரிக் கூறுகள் அமையும். மாதிரி அளவு தெரிவிற்கேற்ப இது அமைந்திருக்கும்.

மாதிரித் தெரிவு முறை

இரண்டு வழிமுறைகள் கையாளப்படுகின்றன. அவை யாவன :

முதல் முறை:

முழுமைத் தொகுதிக் கூறுகள் y_1, y_2, y_3,...........y_N என்றிருப்பின், இதிலிருந்து $k = 3$ என்ற விதிப்படி மாதிரியினை எடுப்பதாகக் கொள்வோம்.

[பொதுவாக, '3-ல் ஒன்று,' '10-ல் ஒன்று.' ....அல்லது k-ல் ஒன்று' என்ற அடிப்படை விதிப்படி கூறுகள் தெரிவு செய்யப் படுவதாகக் கருதுகிறோம்.]

முதல் மூன்று உறுப்புகளினின்று, ஏதேனும் ஒன்றை ராண்டமாகத் தேர்ந்தெடுப்போம். சம வாய்ப்பின் அடிப்படையில் இது செய்யப்படுகிறது. உதராணமாக, y_3 என்ற மதிப்பினைக் கருதுகையில், y_3, y_6, y_9........என்றபடி உறுப்புக்கள், தெரிவு செய்யப்படும் மாதிரியில் இடம்பெறும்.

இதன்படி, N-உறுப்புக்கள் கொண்ட தொகுதியானது, K-உறுப்புகள் கொண்ட n-பிரிவுகளாக ஆக்கப்பட்டு, ஒவ்வொரு பிரிவிலிருந்தும் அல்லது பகுதியிலிருந்தும் ஒவ்வொரு மாதிரிக்கூறு எடுக்கப்படுவதாகக் கொள்ளலாம்.

$$n = \frac{N}{k}$$ என்பது கருதப்படும் விதியாகும்.

அல்லது $k = \frac{N}{n}$ எனவும் கொள்ளலாம். (இது, மாதிரி முறைப் பின்னமாகிய $\frac{n}{N}$ -ன் தலைகீழ் மதிப்பு)

இந்த விதியின்படி, $N = nk$ என்றும் ஆகிறது. சில நிலைகளில், இது ஒரு முழு எண்ணுகவும் அமையும்.

எடுத்துக்காட்டு :

$N = 40$ என்றபடி உள்ள தொகுதியில், $k = 6$ என்று கொண்டு, ஒரு ஒழுங்கு மாதிரியினைத் தெரிவு செய்தல்.

மாதிரி அளவினை

$$n = \frac{N}{k} = \frac{40}{6} = 6{\cdot}67$$ என்ற அளவில் கருதி,

6 அல்லது 7 உறுப்புக்கள் அமைந்த மாதிரிகள் எடுக்க இயலும். மொத்தமாக 6 மாதிரிகள் கிடைக்கும்.

I	II	III	IV	V	VI
y_1	y_2	y_3	y_4	y_5	y_6
y_7	y_8	y_9	y_{10}	y_{11}	y_{12}
y_{13}	y_{14}	y_{15}	y_{16}	y_{17}	y_{18}
y_{19}	y_{20}	y_{21}	y_{22}	y_{23}	y_{24}
y_{25}	y_{26}	y_{27}	y_{28}	y_{29}	y_{30}
y_{31}	y_{32}	y_{33}	y_{34}	y_{35}	y_{36}
y_{37}	y_{38}	y_{39}	y_{40}		

இவற்றில் ஏதேனும் ஒன்றைக் கருதலாம். இது ' 6-ல் ஒன்று , என்ற முறைப்படி எடுக்கப்படுவதாகும். 'ஆறு பேருக்கு ஒருவர்' அல்லது ' ஆறு கூறுகளுக்கு (units) ஒன்று' என்ற நியதியில் $n = 6$ அல்லது 7 அளவுள்ள எந்த மாதிரியினையும் கருதலாம். ஒவ்வொரு மாதிரியும் தெரிவு செய்யப்பட, $\frac{1}{6}$ என்ற சம வாய்ப்பைப் பெறுவதால், இது ராண்டம் மாதிரியாகவும் அமைகிறது.

இரண்டாம் வழிமுறை (Second Method)

இந்த முறையில், எந்த உறுப்பினில் தொடங்கி, மாதிரிக் கூறுகளை (sampling units) எடுப்பது என்பதற்கு ஒரு விதிமுறை வகுக்கப்படுகிறது.

(i) $N = nk$ என்றிருப்பின், துவக்கநிலை (starting point) பின்வருமாறு நிர்ணயிக்கப்படும் :

j என்ற உறுப்பினை அல்லது அலகினை, முழுமைத் தொகுதியிலிருந்து சமவாய்ப்பு அடிப்படையில் தெரிவுசெய்வோம். அதன்பின், $\frac{j}{k}$ ன் மதிப்பை அறிந்து, இதன் மீதமானது எந்த எண்ணாக வருகிறதோ, அந்த எண்ணிற்கேற்ற உறுப்பினைத் துவக்கமாகக் கொள்ளவேண்டும்.

அதாவது $\frac{j}{k} = r$ எனில்,

$r = 1$ ஆனால், y_1 என்பது முதல் கூறாகவும்,

$r = 2$ எனில், y_2 என்பது துவக்கமாகவும்,... இதேபோன்ற முறையில் நிர்ணயிக்க வேண்டும்.

$r = 0$ எனில், k -ஆம் உறுப்பையே (y_k) முதலாவதாகக் கொள்ளவேண்டும்.

உதாரணமாக, $nk = 42$ எனில், $k = 6$ எனக் கொண்டால், $j = 8$ என்ற உறுப்பினை (சமவாய்ப்பு அடிப்படையில்) தெரிவு செய்திடின்,

$$\frac{j}{k} = \frac{8}{6} \text{ என்ற விகித அளவில்,}$$

மீதம் (remainder) வருவது 2 ஆகும்.

அதாவது $r = 2$. $\therefore$ y_2 என்பதைத் துவக்கமாகக் கொண்டு, அதிலிருந்து ஒவ்வொரு 6-ம் உறுப்பினையும் தெரிவு செய்ய வேண்டும். அதாவது, y_2, y_8, y_{14}..............

(ii) $N \neq nk$ என்ற நிலையில், இதே வழிமுறை கையாளப் படுகிறது. மாதிரி அளவுகள் இரண்டு வகைகளில் அமையும். இதில் மாதிரி இடைவெளி (sampling interval), ($N = nq + r'$ என்று கருதி), $r' \leq \frac{n}{2}$ அல்லது $r' > \frac{n}{2}$ என்பதற்கேற்ப q அல்லது $(q + 1)$ என்று எடுக்கப்படலாம்.

உதாரணமாக, $N = 40$, $n = 6$ என்ற நிலையில் $q = 6$, $r' = 9$ (மீதி) என வர, k மதிப்பு 6 எனக் கொள்ளப்படும். 1, 2, 3, 4 எண்கள் உள்ள உறுப்பினுள் ஏதேனும் ஒன்றில் ஆரம்பித்தால் மாதிரி அளவு 7 எனவும் மற்றவற்றில் 6 எனவும் வருவதைக் கண்டோம் $\therefore$ இத்தகைய நிலையில் ஒழுங்குமாதிரி அளவுகள்(systematic samples) ஒரே மாதிரி இருப்பதில்லை. மற்றும், மாதிரிகளின் வாய்ப்பு அளவு $\frac{n}{N}$ என்று அமையும்.

குறியீடு :

y_{ij} என்பது, i-ஆம் ஒழுங்கு மாதிரியில் உள்ள j-ஆம் உறுப்பினை (item or member) குறிக்கும். $j = 1, 2......n$; $i = 1, 2...k$.

$\therefore$ i என்ற மாதிரியின் சராசரி $\bar{y}_i$ ஆகும்.

சராசரி மதிப்பீடு :

தேற்றம் 1 :

ஓர் ஒழுங்கு மாதிரியின் (systematic), மாதிரிச் சராசரியான $\bar{y}_{sy}$ என்பது, முழுமைத் தொகுதிச் சராசரி $\bar{Y}$-ன் பிறழ்ச்சியற்ற மதிப்பீடு ஆகும்.

நிறுவுதல் : $N = nk$ எனக் கொள்க.

$$\bar{y}_i = \frac{1}{n} \sum_j y_{ij} \text{ ஆகும்.}$$

இது i என்ற ஒழுங்கு மாதிரியின் சராசரி.

எனவே,

$$E(\bar{y}_{sys}) = \frac{1}{k}[(\bar{y}_1 + \bar{y}_2 + \dots + \bar{y}_k)]$$

என வரும். ஏனெனில், $\frac{1}{k}$ சமவாய்ப்பு பெறும், k-ஒழுங்கு மாதிரிகளை நாம் தெரிவுசெய்ய முடிகிறது.

$$\therefore \quad E(\bar{y}_{sys}) = \frac{1}{n \cdot k}(y_{11} + y_{12} + \dots + y_{nk})$$

$$= \frac{1}{N}(y_{11} \dots + \dots + y_{nk})$$

$$\therefore \quad E(\bar{y}_{sys}) = \bar{Y}$$

என்று பெறப்படுகிறது.

எடுத்துக்காட்டு 1:

9 பேர்களிடம் உள்ள புத்தகங்கள் வருமாறு :

10, 20. 30, 40, 50, 60, 70, 80, 90.

$k = 3$ என்ற விதியின் மூலம் ஒழுங்கு மாதிரிகள் எடுத்து, சராசரி மதிப்பிடல் :

$N = n.k = 3 \times 3. \quad (n = 3)$

3 மாதிரிகள் எடுக்கப்படலாம்.

அவை

மாதிரி 1 : 10, 40, 70.

2 : 20, 50, 80.

3 : 30, 60, 90.

$\therefore \quad \bar{y}_1 = 40; \quad \bar{y}_2 = 50; \quad \bar{y}_3 = 60.$

எனவே,

$$E(\bar{y}_{sy}) = \frac{1}{3}(40 + 50 + 60) = \frac{150}{3} = 50.$$

மேலும்

$$\bar{Y} = \frac{10 + 20 + \dots + 90}{9} = 50.$$

$\therefore \quad E(\bar{y}_{sys}) = \bar{Y} \qquad \therefore \quad \bar{y}_{sy}$ என்பது $\bar{Y}$-ன் பிறழ்ச்சியற்ற மதிப்பீடாகும்.

இரண்டாம் முறையைப் பயன்படுத்திப் பெறும் சராசரி மதிப்பீடு பிறழ்ச்சியற்றதெனப் பின்வருமாறு விளக்கலாம் :

$k \neq nk$ என்ற நிலை :

$k = 3$ என்று கொண்டு,

$y_1, y_2, y_3, y_4, y_5, y_6, y_7$ என்ற முழுமைத் தொகுதியினின்று ஒழுங்கு மாதிரிகள் எடுப்போம்,

$$N = 7.$$

$$\therefore \quad n = \frac{7}{3} = 2\frac{1}{3}.$$

அதாவது $n = 2$ அல்லது 3 என்றபடி மாதிரிகள் அமையும். அவை :

y_1	y_2	y_3
y_4	y_5	y_6
y_7		

ஆகும். இவற்றின் வாய்ப்புகள் $\frac{3}{7}, \frac{2}{7}, \frac{2}{7}$ ஆகும்.

$$\therefore \quad E(\bar{y}_{sys}) = \left(\frac{3}{7} \cdot \bar{y}_1\right) + \left(\frac{2}{7} \cdot \bar{y}_2\right) + \left(\frac{2}{7} \cdot \bar{y}_3\right)$$

$$= \frac{3}{7}\left[\frac{1}{3}(y_1 + y_4 + y_7)\right] + \frac{2}{7}\left[\frac{1}{2}(y_2 + y_3 + y_5 + y_6)\right]$$

$$= \frac{1}{7}(y_1 + \dots + y_7) = \bar{Y}$$

எனப் பெறப்படும்.

எனவே, பொதுவாக, $\bar{y}_{sys}$ ஆனது $\bar{Y}$ ன் பிறழ்ச்சியற்ற மதிப்பீடாகும்.

சராசரி மதிப்பீட்டின் மாறுபாட்டு அளவு (Variance of $\bar{y}_{sys}$)

இதை அறிவதன்மூலமாக, எத்தகைய சூழ்நிலைகளில், ஒழுங்கு மாதிரிமுறையானது அதிக திட்ப அளவினை, (precision) பெற்று விளங்கும் என்றறிந்து, ராண்டம் முறையினைவிடப் பயனுள்ளதாகக் காணவும் இயலும்.

சில நிபந்தனைகளுடன், படுகை முறைத் தேர்வில் துணைக் கூறுகளை (sub-samples) எடுக்கையில் இம் முறையினைப் பயன்படுத்துவதைப் பின்பு காண்போம்.

வரையறைப்படி,

$$V(\bar{y}_{sys}) = \frac{1}{k}\sum_{k} (\bar{y}_i - \bar{Y})^2$$

ஏனெனில் மொத்தம் k மாதிரிகள் உள்ளன.

மேலும்,

$$V(\bar{y}_{sys}) = \frac{N-1}{N}.\ S^2 - \frac{k(n-1)}{N}.\ S^2_{wsys} \quad \ldots\ldots \quad (1)$$

என்று ஆகும்.

இதில்,

$$S^2_{wsys} = \frac{1}{k(n-1)}.\ \sum_{i}\sum_{j} (y_{ij} - \bar{y}_i)^2$$

என்பது, ஒரே ஒழுங்கு மாதிரியில் அமையும் உறுப்புக்களின் (units) உள்ளீட்டு மாறுபாடு (within).

{ இந்த மாறுபாட்டின், வகுக்கும் காரணி $k(n-1)$ என்பது, மாறுபாட்டுப் பகுப்பாய்வு (analysis of variance) அடிப்படையில் திர்ணயிக்கப்பட்டது. k- மாதிரிகள் மூலம் கட்டுப்பாடற்ற நிலைகள் (degrees of freedom) இவையாகும்.}

மற்றும்

$$S^2 = \frac{1}{N-1}\sum_{i}\sum_{j} (Y_{ij} - \bar{Y})^2$$

{ முழுமைத் தொகுதியின் மாறுபாடு. }

இதன் மூலமாக, அதாவது (1)-ன் மூலம். $V(\bar{y}_{sy})$ ஆனது, முழுமைத் தொகுதி மொத்த மாறுபாட்டு அளவிற்கும், k முறையுடை மாதிரிகளின் மொத்த உள்ளீட்டு மாறுபாட்டு அளவிற்கும் இடையில் உள்ள வேறுபாடாக வரையறுக்கப்படுகிறது.

நிறுவுதல் :

மாறுபாட்டுப் பகுப்பாய்வு விதிமுறைப்படி,

$$(N-1)\,S^2 = \sum_{i}\sum_{j} (y_{ij} - \bar{Y})^2$$

$$= \sum_i \sum_j (y_{ij} - \bar{y}_i + \bar{y}_i - \bar{Y})^2$$

$$= \sum_i \sum_j (y_{ij} - \bar{y}_i)^2 + \sum_i \sum_j (\bar{y}_i - \bar{Y})^2$$

$$= \sum_i \sum_j (y_{ij} - \bar{y}_i)^2 + n \sum_i (\bar{y}_i - \bar{Y})^2$$

ஆனல், $V(\bar{y}_{sys}) = \dfrac{1}{k} \sum_i (\bar{y}_i - \bar{Y})^2$ ஆவதால்,

$$(N-1)\, S^2 = nk.\; V(\bar{y}_{sys}) + k\,(n-1) S_{wsy}^2$$

என வரும்.

எனவே,

$$\frac{(N-1)}{nk} \cdot S^2 - \frac{k\,(n-1)}{nk} \cdot S_{wsys}^2 = V(\bar{y}_{sys})$$

அல்லது

$$V(\bar{y}_{sys}) = \frac{N-1}{N} \cdot S^2 - \frac{k\,(n-1)}{N} \cdot S_{wys}^2$$

என்று பெறுகிறோம்.

துணை முடிவு :

ஒழுங்கு மாதிரிச் சராசரியானது, எளிய சமவாய்ப்பு மாதிரி சராசரியினைவிட அதிக திட்பமுள்ளதாக அமையவேண்டின், $S_{wsy}^2 > S^2$ என அமைவது இன்றியமையாதது.

ஏனெனில்,

$\bar{y}$ என்பது, n-அளவுள்ள SRS மாதிரிச் சராசரியானல்,

$$V(\bar{y}) = \frac{N-n}{N} \cdot \frac{S^2}{n} \text{ என ஆகும்.}$$

மேலும்,

(1)-லிருந்து, $V(\bar{y}_{sy}) < V(\bar{y})$ என்று அமைய வேண்டுமெனில்,

$$\frac{N-1}{N} \cdot S^2 - \frac{k\,(n-1)}{N} \cdot S_{wsy}^2 < \frac{N-n}{N} \cdot \frac{S^2}{n}$$

என்றிருத்தல் அவசியம்.

அதாவது,

$$\frac{k(n-1)}{N} \cdot S_{wsys}^2 > \frac{N-1}{N} \cdot S^2 - \frac{N-n}{N} \cdot S^2$$

அல்லது,

$$k(n-1) \cdot S_{wsys}^2 > \left[(N-1) - \frac{N-n)}{N} \right] \cdot S^2$$

$$= \left[nk - 1 - \frac{nk-n)}{n} \right] \cdot S^2$$

$$= (nk - 1 - k + 1) S^2$$

$$= k(n-1) S^2$$

அதாவது,

$S^2_{wsys} > S^2$ என்றமையவேண்டும்,

இதன்மூலம், ஒரே மாதிரியில் உள்ள கூறுகள் தமமுள் மிக அதிக அளவில் வேறுபடுகையில் (hetero geneous), ஒழுங்கு மாதிரித் தேர்வானது அதிக திட்பமுள்ளதாக அமையும் என்று முடிவு செய்யப்படும்.

ஒழுங்கு மாதிரிகளின் ஓரினத் தன்மை (homogeneity) அல்லது பல்லினத் தன்மையினை (heterogeneity)- அறிய, வேறொரு வகையில், $\bar{y}_{sys}$-யின் மாறுபாட்டை வரையறுக்கிறோம்.

ஒரே ஒழுங்கு மாதிரியின் ஓரினத் தன்மையின் (அதிகம் வேறு பாடுகள் இல்லாத நிலை) அளவினைப் பற்றிய அறிய, அம் மாதிரியில் அமையும் இணைக் கூறுகளின் இடையே (pairs of units) உள்ள உள்ளீட்டு ஒட்டுறவு (intra class) கெழுவினை, அதாவது

$$\rho_w = \frac{E(y_{ij} - \bar{Y})(y_{ij}' - \bar{Y})}{E(y_{ij} - \bar{Y})^2}$$

என்ற கெழுவினைப் பயன்படுத்துகிறோம், ρ_w என்பது, பிரிவு — உள்ளீட்டு ஒட்டுறவுக் கெழு என வரையறுக்கப்படும்.

இதன் மூலமாக,

$$V(\bar{y}_{sys}) = \frac{S^2}{n} \cdot \frac{N-1}{N} \cdot [1 + (n-1)\rho_w].$$

{ρ_w-ன் வரையறையில், பகுதி (Numerator) அளவானது, $\frac{k \cdot n(n-1)}{2}$ வேறு வேறு இணைகளைக் கருதிப் பெறப்பட்ட

சராசரியாகவும், விகுதி அளவானது (Denominator). y_{ij}-யின் எல்லா N மதிப்புகளையும் கருதிப் பெறப்பட்ட சராசரியாகவும் அமைந்திருக்கும். இதை,

$$\rho_w = \frac{2}{(n-1)(N-1)\cdot S^2} \sum_{j=1}^{k} \sum_{j<j'} (y_{ij} - \bar{Y})(y_{ij'} - \bar{Y})$$

என்றும் எழுதலாம்.}

நிறுவுதல்:

வரையறையின்படி,

$$V(\bar{y}_{sys}) = \frac{1}{k} \sum_{i} (\bar{y}_i - \bar{Y})^2$$

இதிலிருந்து,

$$n^2 \cdot k \cdot V(\bar{y}_{sys}) = n^2 \cdot \sum_{i} (\bar{y}_i - \bar{Y})^2$$

$$= \sum_{i=1}^{k} [y_{i1} - \bar{Y}) + \ldots\ldots + (y_{ih} - \bar{Y})]^2$$

என்று பெறலாம்.

வர்க்கமிட்ட உறுப்புக்கள், $\bar{Y}$-லிருந்து பெறப்பட்ட விலக்க வர்க்கங்களின் கூட்டுத் தொகை ஆவதால், அதாவது $(N-1)\,S^2]$,

$$n^2 k . V(\bar{y}_{sy}) = (N-1) . S^2 + 2 \sum_{i} \sum_{j<j'} (y_{ij} - \bar{Y})(y_{ij'} - \bar{Y})$$

$$= (N-1) \cdot S^2 + (n-1)(N-1) \cdot S^2 \cdot \rho_w$$

அல்லது

$$V(\bar{y}_{sy}) = \frac{S^2}{n} \cdot \left(\frac{N-1}{N}\right) \cdot [1 + (n-1)\,\rho_w]$$

என்பது தெளிவாகிறது.

இதிலிருந்து, ஒரே ஒழுங்குமுறை மாதிரியில் உள்ள கூறுகளினிடையில் அல்லது உறுப்புகளினிடையில் உள்ள ஒட்டுறவானது நேர்க்கணியமாக (+ ve Correlation) அமைந்திருப்பின், மாதிரிச் சராசரி மதிப்பீட்டின் மாறுபாடு அதிகரிக்கிறது என்பது புலனாகும். ஒரு சிறு அளவினதான (+) நேர்க்கணிய ஒட்டுறவு கூட அதிக அளவில் இம் மாறுபாட்டினை ப்பாதிக்க இயலுகிறது. ஏனெனில், $(n-1)$ என்பது பெருக்குக் காரணியாக அமைகிறது.

மேற்கண்ட வரையறைகள், $V(\bar{y}_{sys})$ மதிப்பினை S^2 மூலமாக அளிப்பதால், எளிய சமவாய்ப்பு மாதிரியின் மாறுபாட்டுடன் ஒப்பிட இயலுகிறது.

அடுத்து, படுகைச் சமவாய்ப்பு மாதிரி அடிப்படையில் $V(\bar{y}_{sy})$ கொடுக்கப்படும் ஒரு நிலையினைக் காண்போம். இங்கே, படுகைகளாவன : முதல் k உறுப்புகள், இரண்டாவது k உறுப்புக்கள் ...... என்றபடி அமையும். y_{ij} என்பதில் j என்பது படுகை நிலையைக் குறிக்கும். $\bar{y}_{\cdot j}$ என்பது படுகை மாதிரியினைக் குறிப்பிடும்.

தேற்றம் 2 :

$$V(\bar{y}_{sys}) = \frac{S_{wst}^2}{n}\left(\frac{N-n}{N}\right)[1+(n-1)\rho_{wst}]$$

என்பதாகும் இதில்,

$$S_{wst}^2 = \frac{1}{n(k-1)}\sum_{j=1}^{n}\sum_{i=1}^{k}(y_{ij}-\bar{y}_{\cdot j})^2$$

ஆகும். அதாவது, ஒரே படுகையில் அமைந்த உறுப்புக்களின் மாறுபாடு எல்லா n-படுகைகளும் $(k-1)$ என்ற வரையற்ற பாகைகளை (degrees of freedom) அளிப்பதால், $n(k-1)$ என்ற பகுதி காரணி பயன்படுகிறது.

மேலும்,

$$\rho_{wst} = \frac{E(y_{ij}-\bar{y}_{\cdot j})(y_{ij'}-\bar{y}_{\cdot j'})}{E(y_{ij}-\bar{y}_{\cdot j})^2}$$

என்பது, ஒரே ஒழுங்கு மாதிரியில் (Systematic) காணப்படும் இணை உறுப்புகளின், படுகைச் சராசரிகளினின்று பெறும் விலக்கங்களினிடையே உள்ள ஒட்டுறவாகும்.

$$E(y_{ij} - \bar{y}_{\cdot j})^2 = \frac{n(k-1)}{n} \cdot S_{wst}^2$$

(j-ஐப் பொறுத்து E கணக்கிடப்பட்டது)).

$$E(y_{ij} - \bar{y}_{\cdot j})(y_{ij'} - \bar{y}_{\cdot j'})$$

$$= \frac{1}{\frac{n(n-1)}{2}} \cdot \sum_{i} \sum_{j<j'} (y_{ij} - \bar{y}_{\cdot j})(y_{ij'} - \bar{y}_{\cdot j'})$$

$$\therefore \varrho_{wst} = \frac{2}{n(n-1)(k-1)} \cdot \frac{\sum_{i=1}^{k} \sum_{j<j'} (y_{ij} - \bar{y}_{\cdot j})(y_{ij'} - \bar{y}_{\cdot j'})}{S^2_{wst}}$$

என்று அமையும்.

நிரூபணம்:

வரையறைப்படி,

$$n^2 k \cdot V(\bar{y}_{sys}) = n^2 \sum_{i=1}^{k} (\bar{y}_i - \bar{Y})^2$$

$$= \sum_{i} (n\bar{y}_i - n\bar{Y})^2$$

$$= \sum_{1}^{k} [(y_{i1} - \bar{y}_{\cdot 1}) + (y_{i2} - \bar{y}_{\cdot 2}) + \dots + (y_{in} - \bar{y}_{\cdot n})]^2$$

$$\left\{ \because \frac{1}{n} \sum_{j} \bar{y}_{\cdot j} = \bar{Y}. \right.$$

$$= \sum_{i=1}^{k} \sum_{j} (y_{ij} - \bar{y}_{\cdot j})^2$$

$$+ 2 \sum_{i} \sum_{j<v} (y_{ij} - \bar{y}_{\cdot j})(y_{ij} - \bar{y}_{\cdot j})$$

$$= n(k-1) \cdot S_{wst}^2 + n(n-1)(k-1)\, S_{wst}^2\, \varrho_{wst}$$

$$\therefore\ V(\bar{y}_{sys}) = \frac{n(k-1)}{n^2k} \cdot S_{wst}^2 + \frac{n(n-1)(k-1)}{n^2k} S_{wst}^2 \cdot \varrho_{wst}$$

$$= \frac{(nk-n)}{n^2k} \cdot S_{wst}^2 + \frac{(n-1)(nk-n)}{n^2k} \cdot S_{wst}^2 \cdot \varrho_{wst}$$

$$= \frac{(nk-n)}{n^2k} \cdot S_{wst}^2\, (1 + \overline{n-1} \cdot \varrho_{swt})$$

$$\therefore\ V(\bar{y}_{sys}) = \frac{N-n}{N} \cdot \frac{S_{wst}^2}{n} [1 + (n-1)\, \varrho_{swt}]$$

என்று பெறப்படும்,

துணை முடிவு:

$\varrho_{wst} = 0$ என்று அமையும் நிலையில், முறையுடை மாதிரியானது, ஒவ்வொரு படுகையிலும் ஓர் உறுப்பினைக் கொண்ட படுகை சமவாய்ப்பு மாதிரியின் அளவு திட்பம் உள்ளதாக இருக்கும்.

∴ இத்தகைய படுகை முறையில்.

$$V(\bar{y}_{sy}) = \frac{N-n}{N} \cdot \frac{S^2_{wst}}{n} \text{ ஆகும்.}$$

எடுத்துக் காட்டு 1

$N = 9$ என்ற முழுமைத் தொகுதி :

1, 2, 3, 4, 5, 6, 7, 8, 9.

$n = 3$ என்ற அளவில் ஒழுங்கு மாதிரிகளை எடுத்து, $\bar{y}_{sy}$ கணக்கிட்டு, $E(\bar{y}_{sy}) = \bar{Y}$ என்று காட்டுக.

$N = nk \quad \therefore\ 9 = 3k$ அல்லது, $k = 3$,

∴ மூன்று மாதிரித் தேர்வுகள் அமையக் காணலாம்.

மாதிரி 1	மாதிரி 2	மாதிரி 3
1	2	3
4	5	6
7	8	9
y_1 : 12	y_2 : 15	$\bar{y}_3$: 18

$$\bar{y}_1 = \frac{y_1}{n} = \frac{12}{3} = 4\ ;$$

$$\bar{y}_2 = \frac{15}{3} = 5\ ;$$

$$\bar{y}_3 = \frac{18}{3} = 6.$$

மேலும்,

$$E(\bar{y}_{sy}) = \frac{1}{3}(4+5+6)$$

$$= \frac{15}{3} = 5$$

$$\bar{Y} = \frac{1+2+3+\ldots\ldots+9}{9} = \frac{45}{9} = 5$$

$\therefore\ E(\bar{y}) = \bar{Y}$ என்பது தெளிவு.

எடுத்துக் காட்டு : 2

126 மாணவர்களின் ஒரு வாரக் கட்டணச் செலவுகள் தரப்படுகின்றன. இவற்றிலிருந்து 7 அளவு கொண்ட ஒழுங்கு மாதிரியினைத் தெரிந்தெடுத்து, சராசரியும், மாறுபாட்டு அளவும் (பரவற்படி) காண்க : (ரூபாயில்) :

36	70	20	35	37	20	62	36
54	35	25	36	65	24	33	48
18	30	8	29	39	9	44	12
24	19	39	3	27	8	42	14
31	31	12	33	15	30	22	7
44	24	54	8	30	22	66	30
6	23	39	18	36	22	14	32
41	9	30	30	35	22	54	26
48	25	39	60	47	29	54	36
29	52	24	48	27	36	42	42
18	33	52	25	34	36		

45	25	48	32	
48	7	32	12	
38	24	3	30	
24	7	31	26	
42	23	31	16	
28	24	27	30	
24	27	3	28	
26	60	30	60	
27	60	29	36	
25	62	41	51	(B. sc., Sep '75)

$N = 126$; $n = 7$

$$\therefore \quad k = \frac{126}{7} = 18$$

18 உறுப்புகள் கொண்ட 7 பகுதிகளாகத் தொகுதியானது பிரிக்கப்பட்டு, ஏதேனும் ஓர் எண்ணில் ராண்டமாகத் துவங்கி, ஒவ்வொரு 18 வது உறுப்பினையும் அல்லது மதிப்பையும் தெரிவு செய்ய வேண்டும்.

20-ஆவது உறுப்பினைக் கருதுக.

இதற்கு $\frac{20}{k} = \frac{20}{18}$ என்று பெற்று, மீதியானது $r = 2$ ஆகும். எனவே, 2-வது உறுப்பு அல்லது மதிப்பானது முதலில் தெரிவு செய்யப்பட்டு, அதிலிருந்து ஒவ்வொரு 18-ஆம் உறுப்பும் அல்லது மதிப்பும் தேர்ந்தெடுக்கப்படும்.

அதாவது 2, 20, 38, 56, 74, 92, 110, என்ற ஏழு கூறுகள் (units) தெரிவு செய்யப்படுகின்றன. அவையாவன : (மதிப்புக்கள்).

ரூ 54, 25, 33, 20, 54, 28, 31
என்பவை ஆகும்.

இந்த ஏழு உறுப்புகளும், நமக்குத் தேவைப்படும் மாதிரியைத் தரும். ($n = 7$)

இதன் சராசரி :

$$\bar{y} = \frac{245}{7} = 35$$

$\therefore$ சராசரியாக ரூ. 35 செலவாகிறதெனலாம்.

மாறுபாட்டு அளவு :

அதாவது $V\ (y)$.

$$V\ (y) = S^2 = \frac{1}{n-1} \Sigma (y_i - \bar{y})^2$$

$$= \frac{1}{6} [(19)^2 + (-10)^2 + (-2)^2 + (-15)^2 + (19)^2 + (-7)^2 + (-4)^2]$$

$$= \frac{1}{6} [361 + 100 + 4 + 225 + 361 + 49 + 16]$$

$$= \frac{1116}{6}$$

$\therefore\ S^2 = 186$ ரூ.

என்று காணப்படும்.

எடுத்துக் காட்டு 3

முழுமைத் தொகுதி :

1, 2, 3, 4, 5, 6, 7, 8, 9.

$k = 3$ என்ற நியதியில் பெறும் ஒழுங்கு மாதிரிகள்

I	1	4	7
II	2	5	8
III	3	6	9.

$V\ (\bar{y}_{sy})$-யினை மதிப்பிடல் ;

ஒவ்வொரு மாதிரியிலும் உள்ள உறுப்புகள் வேறுபாட்டுடன் அமைவதைக் காணலாம்.

முழுமைத் தொகுதி சராசரி :

$$\bar{Y} = \frac{\Sigma \Sigma y_{ij}}{N}$$

$$= \frac{1 + 2 + \ldots\ldots + 8 + 9}{9} = 5.$$

$$S^2 = \frac{1}{N-1} \Sigma \Sigma (y_{ij} - \bar{Y})^2 = \frac{(-4)^2 + \ldots\ldots + (4)^2}{(9-1)}$$

$$= \frac{16 + \ldots\ldots + 16}{8} = \frac{60}{8} = 7{\cdot}5$$

பிரிவு—உள்வீட்டு ஒட்டுறவுக் கெழு :

$$\xi = \frac{2 \sum\limits_{i} \sum\limits_{j<j'} (y_{ij} - \bar{Y})(y_{ij'} - \bar{Y})}{(n-1) \cdot (N-1) \cdot S^2}$$

$$= \frac{2}{(2 \times 60)} [y_{11} - \bar{Y})(y_{12} - \bar{Y}) + \ldots\ldots + \ldots\ldots]$$

$$= \frac{1}{60} [(1-5)(4-5) + (1-5)(7-5) + (4-5)(7-5) + \ldots\ldots]$$

$$= \frac{1}{60} [(-6-9-6] = \frac{-21}{60} = \frac{-7}{20}.$$

எனவே,

$$V(\bar{y}_{sy}) = \frac{(N-1) \cdot S^2}{N \cdot n} [1 + (n-1)\,\xi]$$

$$= \frac{8}{(9 \times 3)} \cdot \frac{60}{8} \left[1 + 2 \cdot \left(\frac{-7}{20}\right)\right]$$

$$\therefore \quad V(\bar{y}_{sy}) = \frac{2}{3} = 0{\cdot}67,$$

பயிற்சி

இதே தொகுதியை, ஒரு படித்தான பகுதிகளாகப் பிரித்து அதன்மூலம் மாதிரிகள் எடுத்து, மேற்கண்டவற்றைக் காண்க.

(குறிப்பு : 1 2 3 | 4 5 6 | 7 8 9 எனப் பிரிக்க.)

சில நிலைகள் (ஒழுங்கு மாதிரியில்).

முழுமைத் தொகுதியின் இயல்பைப் பொறுத்தே, ஒழுங்கு முறை மாதிரித் தேர்வு, படுகை முறைத் தேர்வைவிட சிறந்ததா என்று முடிவிடமுடியும். சில தொகுதிகட்கு, மற்றும் சில உ அளவுகட்கு, $V(\bar{y}_{sy})$ பெரும் அளவில் அதிகரிக்கவும் நேரிடும்.

ஏாண்டம் வரிசையில் (Random order) தொகுதி உறுப்புகட்கு (units) வரிசை எண்கள் அளிக்கப்படின், (ordering), ஒழுங்கு முறை மாதிரி எதிர் நோக்கும் திட்டம் பெற்றிருக்கலாம். உதாரணமாக, அகர வரிசையில் (alphabetical) அமைந்த வரிசை முறையானது, கருதப்படும் இயல்புகள் ;ஒன்றோடொன்று

தொடர்பற்றும், ஒரே மாதிரிப் போக்கு (trend) இல்லாமலும் அமைந்திடக்காணலாம். இந்நிலையில், ஒழுங்கு முறையானது, சமவாய்ப்பு முறையைப் போலவே இருக்கும், என எதிர் பார்க்கலாம். ஆனால், எல்லாத் துணைத் தொகுதிகளையும் கருதும் போதுதான், (அதாவது $N!$),

$E(V_{sy}) = V_{ran}$. என்று பெற இயலும். பொதுவாக, நடை முறையில், பல வரம்புடைத் தொகுதிகளின் தொடர் பற்றியே மாதிரி முறையானது பயன்படுத்தப்படுகிறது என்பது குறப்பிடத் தக்கது.

வரையறை

μ என்ற சராசரி மதிப்புடன் அமைந்த ஒரு பெரிய தொகுதியினின்று $y_1 \ldots y_N$ என்ற மதிப்புகளை வாய்ப்பின் அடிப்படையில் தெரிவு செய்கையில் (இதுவும் ஒரு வரம்புடை தொகுதியாகும்),

$E'(V_{sy}) = E'(V_{ran})$ என்பது உண்மையாகும். எல்லா y_i-ம் μ என்ற ஒரே சராசரியும், மாறக்கூடிய மாறுபாடும் பெற்றிருத்தல் வேண்டும். (கொக்ரான்—Cochran).

[E' என்பது பல்வேறு தொகுதிகளின் அடிப்படையில் சராசரி மதிப்பிடுவதைக் குறிக்கும்.]

இதில் $E'(y_i - \mu)(y_j - \mu) = 0 \quad (i \neq j)$;

$$E'(y_i - \mu)^2 = \sigma_i^2.$$

[இதன் நிருபணம் தேவையில்லை].

நேர்கோட்டுப் போக்குடன் (linear trend) அமைந்த தொகுதிகள்:

முழுவதும் நேர்கோட்டுப் போக்கில் அமைந்த தொகுதியைக் கருதுகையில், முடிவுகளை எளிதில் பெறமுடிகிறது. இதில், V_{sy} மற்றும் V_{st} ஆகிய இரண்டும் V_{ran} மதிப்பைவிடக் குறைந்தவை போல காணப்படும். (V_{st} என்பது, ஒவ்வொரு படுகையிலும் ஒரே ஒரு உறுப்புக் கொண்டதாக கருதப்படும்).

மேலும், $V_{sy} > V_{st}$,

ஏனெனில், ஒரு படுகையில் எடுக்கப்பட்ட ஒழுங்கு மாதிரி மிகச் சிறிய அளவில் இருப்பின், அவ்வாறே ஏனைய படுகைகளிலும் இருக்கும்.

$y_i = i$ என்று கொள்க.

$$\sum_N i = \frac{N(N+1)}{2} \; ; \; \sum_N i^2 = \frac{N(N+1)(2N+1)}{6}$$

எனவே,

$$S^2 = \frac{1}{N-1}[\Sigma y_i^2 - (N\bar{Y}^2)]$$

$$= \frac{1}{N-1}\left[\frac{N(N+1)(2N+1)}{6} - \frac{N(N+1)^2}{4}\right]$$

$$\therefore \quad S^2 = \frac{N(N+1)}{12} \qquad \ldots\ldots \quad (1)$$

எனவே,

$$V_{ran} = V(\bar{y}_{ran})$$

$$= \frac{N-n}{N} \cdot \frac{S^2}{n}$$

$$= \frac{n(k-1)}{N} \cdot \frac{N(N+1)}{12n}$$

$$\therefore \quad V_{ran} = \frac{(k-1)(N+1)}{12} \qquad \ldots\ldots \quad (2)$$

S_{wsy}^2 காண்பதற்கு, (1)-ல் N-க்கு k என்று பிரதியிட,

$$V_{st} = \frac{N-n}{N} \cdot \frac{S_{wsy}^2}{n} = \frac{n(k-1)}{nk} \cdot \frac{k(k+1)}{12n}$$

$$= \frac{(k^2-1)}{12n} \qquad \ldots\ldots \quad (3)$$

மேலும் $\sum_k (\bar{y}_i - \bar{Y})^2 = \frac{k(k^2-1)}{12}$ எனப் பெறப்படும்.

[∵, ஒழுங்கு முறையில், இரண்டாம் மாதிரி சராசரி முதலாவதைவிட '1' அதிகரித்தும், இரண்டாவதைவிட மூன்றாவது மாதிரி சராசரி '1' கூடவும், ...இருக்கும். எனவே, $\bar{y}_i$ ஆனது, 1. 2 ...... k என்று பெறப்படும்.]

இதன்மூலம் $V_{sy} = \frac{1}{k} \sum_k (\bar{y}_i - \bar{Y})^2 = \frac{k^2-1}{12}$ ஆகும்.

...... (4)

எனவே, (2), (3), (4) இவற்றை ஒப்பிட்டு,

$$V_{st} = \frac{k^2-1}{12n} \leq V_{sy} = \frac{k^2-1}{12} \leq V_{ran}$$

$$= \frac{(k-1)(N+1)}{12}$$

என்று பெறலாம், $n = 1$ எனில் இவை மூன்றும் சமமாகும்.

∴ முடிவு: நேர்கோட்டுப் போக்குத் தன்மையை நீக்க, ஒழுங்கு முறையானது *SRS* முறையைவிட அதிகத் திறனுள்ள தாய் அமையும். ஆனால், இது படுகை முறையினைவிடக் குறைந்த திறனுடையதாகும்.

எல்லைகளின் திருத்தம் (End Correction):

நேர்கோட்டுப் போக்குடன் தொகுதியானது அமைந்திடின், $\bar{Y}$-ன் மதிப்பீடானது பின்வரும் எடைகளைப் பயன்படுத்துவதன் மூலம், சிறந்ததாக்கப் படலாம், அதாவது, மாதிரியின் முதலாவது மற்றும் இறுதிக் கூறுகட்கு, முறையே

$$\frac{1}{n} + \frac{(2r-k-1)}{2(n-1)\cdot k}\ ;\ \frac{1}{n} - \frac{(2r-k-1)}{2(n-1)k}$$

என்ற எடைகளை உபயோகிக்கவேண்டும். இது, $Y_i = a + bi$ $(i = 1, 2 \ldots\ldots N)$ என்ற பொதுவான நேர்கோட்டு விதிப்படி அமையும் தொகுதியில் பயன்படுத்தப்படுகையில் மாறுபாடானது பூச்சியமாக அமையத்தக்க விதத்தில் நிர்ணயிக்கப்படுவதாகும்.

முதல் மற்றும் கடைசி உறுப்புகள், $\left(\frac{1}{n} + x\right)$, $\left(\frac{1}{n} - x\right)$ என்று

ஆனால், r-ஆவது ஒழுங்கு மாதிரியில், மதிப்பீட்டினை

$$\bar{y}_r' = \frac{1}{n} \sum_{j=0}^{n-1} \{ a + b(r + kj) + x(a + br) - x[a + b(r + \overline{n-1}\cdot k)] \}$$

$$= a + b\left(r + \frac{n-1}{2}\right) - x \cdot b(n-1)k$$

என்று பெற்று, அதனைத் தொகுதிச் சராசரியான

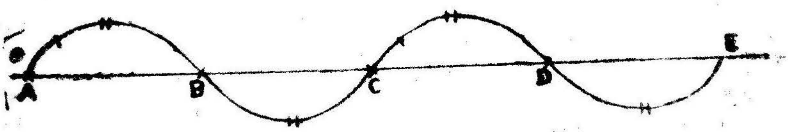
$a + \frac{1}{2} b(N+1)$-க்குச் சமமானதாக்கி, x-க்குத் தீர்வு காண

$x = \dfrac{2r-k-1}{2(n-1)k}$ என்று பெறப்படும்.

இவை, ஓர மதிப்புகளின் அல்லது கடைசித் திருத்தம் (end correc tions) எனப்படும். இவற்றின்மூலம், மாறுபாடானது (Variance) குறைக்கப்படுகிறது.

காலச் சுழற்சி (Periodic Variation) ஒழுங்குள்ள தொகுதிகள் :

ஏதேனும் ஒரு தொகுதியில் (population) சிறு அளவுள்ள அல்லது நடுத்தர அளவுள்ள அல்லது பெரிய அளவுள்ள மதிப்புக்கள் சுழன்று வரும் அமைப்பும் இருக்கப் பெறின், மிகுந்த கவனத்துடன் ஒழுங்கு மாதிரி முறையானது பயன்படுத்தப்பட வேண்டும். கூறுகளின் மதிப்புகள் பின்வரும் ஒழுங்குப்படி அமைந்திருப்பதாகக் கொள்வோம்:



$AB = BC = CD = DE$. AC (அல்லது CE) என்பது ஒரு சுழற்சியின் கால அளவு (காலவட்ட சுழற்சி) (period) எனப்படும். மாதிரித் தேர்வின் இடைவெளியானது சமமாக (அல்லது இக் கால அளவின் பெருக்கற் பலனாக) இருப்பின், எடுக்கப்படும் மாதிரியில் அமையும் கூறுகள் ஒரே மதிப்பினதாக அமைகையில், ஒழுங்குமுறை மதிப்பீடுகள் அதிக அளவில் வேறுபடுகின்றன. ('ı' குறியிடப்பட்ட பகுதிகளைக் காண்க). இவ் வளைவின் கால அளவு மதிப்பை (periodicity) அறிந்திடின், மாதிரிக் கூறுகளின் இடைவெளியினை, கால அளவின் பாதி மதிப்பினது ஒற்றைப் பெருக்கற் பலனாகக் கொள்ளலாம் (odd multiple). அப்போது, மாறுபாடும் அளவு (மதிப்பீடுகட்கு) மிகவும் குறைக்கப்படும். ('||' என்ற குறியிடப்பட்ட இடங்கள்).

ஒரு தொகுதியானது ஏதேனும் ஒரு பொதுவான இயல்பின் அடிப்படையில் சம அளவுக் கூறுகள் உள்ள பிரிவுகளாகப் பகுக்கப்படுகையில், ஒழுங்கு முறை மாதிரியில் ஏறக்குறைய ஒரே தன்மையுள்ள கூறுகளே தெரிவுபெற வாய்ப்புண்டு. மக்கட் கணிப்பில்இது அதிக அளவில் நிகழ்கிறது. உதாரணமாக, இணையான தெருக்களைப் பிரிவுகளாகக்கருதி, ஒழுங்குமுறையில் மாதிரித் தெரிவு செய்கையில், செல்வந்தர்கள் முதல் வரிசையில் (முக்கிய சாலைக்கு அருகில்) வசிக்கவும், வசதி குறைந்தோர் அதன் பின்னரும் இருக்கும்படியான சூழ் நிலையில், ஒழுங்கு அடிப்படையானது அதிக பிறழ்ச்சியுள்ள தேர்வுக்கு வழிவகுக்கும். எனவே, தக்க அமைப்பின் மூலமாக, இத்தகைய 'திரும்பத் திரும்ப' வரும் சுழற்சியினைத் தவிர்ப்பது இன்றியமையாதது.

உள்ளிட்ட ஒட்டுறவு அமைந்த முழுமைத் தொகுதிகள்: (Autocorrelated Populations)

கொடுக்கப்படும் இயற்கையான முழுமைத் தொகுதிகளில், i, j என்ற உறுப்புகள் அருகருகே இருப்பின், y_i, y_j ஆகியவை ஒரே மாதிரி இருப்பதற்கு அதிக வாய்ப்பு உண்டு. அந் நிலையில், y_i, y_j இரண்டும் நேர்க்கணிய உடன் தொடர்பு (+vely correlated) உள்ளவையாக இருக்கும். $i-j$ இடைவெளியானது (distance) அதிகரிக்க அதிகரிக்க, இந்த ஒட்டுறவு (அல்லது உடன் தொடர்பு) மிகவும் குறைய நேரிடும்.

இதை அறிவதற்கு, u என்ற அளவில் இடைவெளியோடு அமைந்த உறுப்புகளின் இடைப்பட்ட ஒட்டுறவு மதிப்புகளைக் (Correlations) கண்டு, அவற்றைத் தக்கதொரு விளக்கப்படத்தில் (correlogram) குறிக்கலாம். ஆனால், வரம்புற்ற முழுமைத் தொகுதியின் காரணமாக, சில ஒழுங்கீனப் போக்குகள் (Irregularites) காணப்படும். பல்வேறு முழுமைத் தொகுதிகளை ஒரு மிகப் பெரிய தொகுதியிலிருந்து (super-population) எடுத்து, அந்த அடிப்படையில் ஒழுங்கு முறை மற்றும் படுகைமுறை இரண்டிற்கும் உள்ள ஒப்புமை (comparison) அறியப்படலாம்.

y_i $(i: 1, 2 \ldots N)$ என்ற தொகுதியானது, $E'(y) = \mu$; $E'(y_i - \mu)^2 = \sigma^2$; $E'(y_i - \mu)(y_{i+u} - \mu) = \xi_u \sigma^2$ என்ற இயல்புகள் அமைந்த மிகப் பெரிய தொகுதியினின்று கருதப்படுவதாகக் கொள்கிறோம்.

மேலும், $u < v$ எனில், $\xi_v \geq \xi_u \geq 0$ ஆகும். ξ_u என்பது அத் தொகுதிகளில் உள்ள உறுப்புகளின் இடையில் u இடைவெளி இருக்குமாறு அமைந்த இணைகளின் ஒட்டுறவு.

ஒழுங்கு மாதிரியில், சராசரி மாறுபாட்டு மதிப்பானது

$$E' V_{sys} = E' [E(\bar{y}_{sys} - \bar{Y})^2].$$

இத்தகைய முழுமைத் தொகுதிகட்குப் படுகைமுறை மாதிரித் தேர்வு சிறந்ததாக அமைகிறது.

k-யின் சில மதிப்புகட்கு, சமவாய்ப்பு முறையைவிட ஒழுங்கு முறை மாதிரித் தேர்வானது சிறந்த முடிவுகளை அளிக்காத நிலைகளும் உண்டு.

உள்வீட்டு ஒட்டுறவுக் கெழுக்களுக்கு வரைந்த வளைவானது (correlogram), மேல்நோக்கிக் குவிவாக (concave upwards) அமைந்திருப்பதாகக் கொண்டு, பின்வரும் உண்மையினைக் காண்போம்:

தேற்றம்

கொடுக்கப்பட்ட நிபந்தனைகளுடன்,

$$\xi_i^2 = \xi_{i+1}\ \xi_{i-1} + 2\xi_i \geq 0 \qquad [i : 2, 3, \ldots (kn-2)]$$

என்றும் இருப்பின்,

$$E'V_{sys} \leq E'V_{st} \leq E'V_{ran}$$

என்ற தொடர்பு நிறுவப்படும்.

மேலும், $E'(y_i) = \mu_i$; $E'(y_i - \mu_i)^2 = \sigma^2$ என்று இருந்தால், மூன்று சராசரி மாறுபாடுகளும் ஒரே அளவில் அதிகரிக்கப்படுவதால், சமனிலிகள் (inequalities) அமையும். அதாவது $E'V_{sys} < E'V_{st} < E'V_{ran}$

பயிற்சி

1. 200 அடி நீளப் பரப்பில் ஒவ்வொரு பிரிவிலும் அமைக்கப் பட்ட நாற்றுக்களின் எண்ணிக்கைகள் வருமாறு :

1 — 20	21 — 40	41 — 60	61 — 80	81 —100	101 - 120	121—140	141—160	161—180	181—200
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
8	20	26	34	31	24	18	16	36	10
6	19	26	21	23	19	13	12	8	35
6	25	10	27	41	28	7	8	29	7
23	11	41	25	18	18	9	10	33	9
25	31	30	32	15	29	11	12	14	12
16	26	55	43	21	24	20	20	13	7
28	29	34	33	8	33	16	17	18	6
21	19	56	45	22	37	9	12	20	14
22	17	39	23	11	32	14	7	13	12
18	28	41	27	3	26	15	17	24	15
26	16	27	37	4	36	20	21	29	18
28	9	20	14	5	20	21	26	18	4
11	22	25	14	11	43	15	16	16	4
16	26	39	24	9	27	14	18	20	9
7	17	24	18	25	20	13	11	6	8
22	39	25	17	16	21	9	19	15	8
44	21	18	14	13	18	25	27	4	9
26	14	44	38	22	19	17	29	8	10
31	40	55	36	18	24	7	31	8	5
26	30	39	29	9	30	30	29	10	3

ஒவ்வொரு 20 வது அடியையும் கருதிய ஒழுங்கு மாதிரியின் சராசரி, மற்றும் அதன் மாறுபாட்டு மதிப்பீடு இவற்றைப் பெறுக. அதாவது, $\bar{y}_{st}$, மற்றும் $V(\bar{y}_{st})$ இவற்றைக் கணக்கிடுக.

இந்த மாறுபாட்டு மதிப்பீட்டினை (i) எளிய சமவாய்ப்பு முறை (ii) 2 உறுப்புகளை ஒவ்வொரு படுகையிலும் கொண்ட படுகை மாதிரி முறை இவற்றின்மூலம் பெறும் $n = 10$ என்ற அளவுள்ள மாதிரிகளிலிருந்து பெற்ற மாறுபாட்டு அளவுடன் $V(\bar{y}_{st})$) ஒப்பிடுக.

[$\Sigma (y_i - \bar{Y})^2 = 23601$ என்று கொள்க.]

2. 290 விளை நிலங்களிலிருந்து எடுக்கப்படும் 29 மாதிரிக் கூறுகள் உள்ள ஒழுங்கு மாதிரியில் பின்வரும் பயிரிடும் நிலப் பரப்பு கொடுக்கப்படுகிறது. இந்த விவரங்களைப் பயன்படுத்தி,

(i) மொத்த விளைநிலப் பரப்பு Y யினை மதிப்பீடு செய்க.

(ii) $\hat{Y}$ ன் மாறுபாட்டு அளவினைக் கணிக்க (அதாவது $V(\hat{Y})$].

நிலம்	y	நிலம்	y	நிலம்	y
1	0·0	11	2·8	21	2·3
2	0·9	12	2·6	22	2·9
3	0·0	13	2·3	23	2·1
4	0·0	14	3·5	24	6·3
5	0·3	15	2·4	25	8·2
6	0·1	16	3·8	26	5·4
7	0·5	17	4·1	27	6·5
8	3·1	18	4·9	28	6·6
9	2·8	19	6·0	29	4·1
10	2·7	20	5·4		

(y : ஹெக்டேரில்)

3. 2, 9, 5, 4, 8   6, 6, 4, 7, 2   4, 9, 4, 8, 3
3, 6, 9, 9, 2   2, 4, 9, 1, 5   1, 3, 5

என்ற தொகுதியிலிருந்து $k = 6$ என்றபடி, மாதிரிகளைத் தெரிவு செய்து, $Y, \bar{Y}$ ஆகியவற்றை மதிப்பீடு செய்து $E(\bar{y}_{sy}) = \bar{Y}$ என்று காட்டுக.

மேலும், $\bar{y}_{sy}$-ன் மாறுபாட்டினைக் காண்க.

4. 1200 பேர்களுள்ள ஒரு கல்லூரியினின்று, $n = 200$ என்ற ஒழுங்கு மாதிரியினைத் தெரிவு செய்யும் வழிமுறையினை விளக்குக.

5. முழுத்தொகுதியானது
1, 2, 3, 5, 6, 7, 8, 9, 10, 11, 12

எனில்,

(i) $k = 4$ என்றபடி ஒழுங்கு மாதிரிகள் எடுத்து, ξ-வின் மதிப்பினைக் கணக்கிட்டு, $V(\bar{y}_{sy})$ யினை மதிப்பிடுக.

(ii) ஒழுங்கு மாதிரி முறையினை, சமவாய்ப்பு முறையுடன் திறனளவின் அடிப்படையில் ஒப்பிடுக.

6. 1, 4, 7, 10, 2, 5, 8, 11, 3, 6, 9, 12
என்ற தொகுதியிலிருந்து '4-ல் ஒன்று' என்று வரும் எல்லா ஒழுங்கு மாதிரிகளையும் தெரிவு செய்து, ξ-வினை மதிப்பிடுக.
அதன் மூலமாக $V(\bar{y}_{sy})$ யினைக் கணக்கிடுக.
சமவாய்ப்பு முறையுடன் ஒப்பீடு செய்க.

7. திரள் மாதிரிமுறை

(Cluster Sampling)

I இரு கட்ட மாதிரி முறை :

[ஒரு திரளின் மாதிரி — 'm' திரள்களை தேர்ந்தெடுத்தல் — சராசரி மதிப்பீடு — $\hat{Y}$-ன் மாறுபாடும், அதன் மாதிரி மதிப்பீடும்— $V(\hat{Y}_c)$-யின் மதிப்பீட்டு அளவை—திரள்கள் அளவும், துணைக்கூறு அளவும் சமமான நிலை. எடுத்துக்காட்டுகள்; பயிற்சிகள்—திரள்முறை. எளிய ராண்டம் முறை—ஒப்புமை. — திரள் முறையில், மாதிரிப் பங்கீடு.]

II திரள் முறையும் வீத மதிப்பீடும்

[P-யின் மதிப்பீடு — $\hat{P}$-ன் மாறுபாடு — $V(\hat{P})$-ன் மாதிரி மதிப்பீடு—எடுத்துக் காட்டுகள், பயிற்சிகள்]

III பல கட்ட மாதிரி முறை

IV பல்தோற்ற மாதிரி முறை

இருகட்ட மாதிரிமுறை (Two Stage Sampling Method)

சில ஆய்வுகளில், மாதிரி எடுக்கப் பயன்படுத்தப் பெறும் கூறுகள் (units), சிறு சிறு கூறுகளின் தொகுப்புகளாக அல்லது 'திரள்'களாக (clusters) அமைந்திருக்க நேரிடும். பெரும்பாலான நிலைகளில், முழுமைத் தொகுதிக் கூறுகள் பற்றிய விரிவான பட்டியல் (list) அமைக்கப்படாமலும், அத்தகைய பட்டியலை அமைப்பது அதிகச் செலவுடையதாகவும் காணப்படும். வெவ்வேறு நிலப்பரப்புகளில் வசிக்கும் மக்கள், இருக்கும் வீடுகள், கால்நடைகள் போன்ற பலவற்றையும்பற்றித் தெளிவான விவரப் பட்டியல் கிடைப்பதில்லை. எனவே, பல்வேறு தொகுதிகளாகப் (blocks) பிரித்து, அவற்றிலிருந்து மாதிரி எடுக்கப்படும். அதன் பின், ஏதேனும் ஒரு நியதியின் அடிப்படையில் தெரிந்தெடுக்கப்படும் திரள்களின் உட்கூறுகளினின்று மாதிரித் தேர்வு நடைபெறலாம். இது நடைமுறைக்கு எளிதாகவும், குறைந்த செலவுள்ளதாகவும், முடிவுகள் எடுப்பதில் காலதாமதம் ஏற்படாததாகவும் உள்ள ஒரு வழிமுறையாகிறது.

ஒழுங்கு மாதிரிமுறையில் (Systematic samp ing), ஒரு முழுத் தொகுதியைப் பல பிரிவுகளாகப் பகுத்து, ஒவ்வொரு பிரிவிலும் உள்ள கூறுகள் ஒரு ஒழுங்கான இடைவெளியில் பரவலாக அமையும் நிலையில், அவற்றிலிருந்து ஏதேனும் ஒரு பகுதியானது தெரிவு செய்யப்பட வாய்ப்பு அமைந்திடும் தன்மையைக் காணலாம்.

சில குடும்பங்களை, அல்லது சில வீடுகளை, அல்லது சில பண்ணைகளை ஒன்றாக இணைத்து ஒரு திரளானது உருவாக்கப் படுகிறது. இத் திரள்களினின்று மாதிரி எடுக்கப்படுவது, அதன் மாறுபாட்டினை மிகவும் பாதிப்பதைப் பின்பு விரிவாகக் காண் போம். அடுத்தடுத்து உள்ள கூறுகளை இணைப்பது நடை முறைக்கு எளிது. அடுத்தடுத்த இரண்டு அல்லது அதற்குமேல் உள்ள குடும்பங்களை ஒன்றாக இணைத்து ஒரு திரள் அல்லது தொகுப்பு உருவாக்கப்படலாம். அல்லது, ஒரு தொழிலகத்தில் பல்வேறு கால இடைவெளிகளில் உற்பத்தியாகும் அளவினைக் கொண்டு, திரள்கள் அமைக்கப்பெறலாம். இத் திரள்கள் கருதப்படுகையில், ஏதேனும் ஒன்றன் தேர்வு மற்றொன்றின் தேர்வினை விலக்கக்கூடியதாய் (mutually exclusive) அமைவது நன்று.

(i) ஒரு திரளின் மாதிரி (Sampling of One Cluster)

வரம்புடைய ஒரு முழுமைத் தொகுதியானது, ஒவ்வொன்றும் N கூறுகள் கொண்ட ஒன்றை மற்றொன்று விலக்கக் கூடியதாக (mutually exclusive) அமைந்த M திரள்களாகப் பிரிக்கப்படு கிறது. இந்த M திரள்களில், ஒன்று தெரிவு செய்யப்படுகிறது. y_{ij} என்ற மதிப்பு i-ஆம் திரளில் உள்ள j-ஆம் உறுப்பின் மதிப்பாகக் கொண்டு,

$$\bar{Y}_i = \frac{1}{N} \sum_{j} Y_{ij} \text{ என்று அறியலாம்.}$$

இதன்மூலம்,

$$\bar{Y} = \frac{1}{M} \sum_{i=1}^{M} \bar{Y}_i \text{ என்பது பெறப்படும்.}$$

எனவே, மாதிரித் திரளின் சராசரியானது, $\bar{Y}$-ன் பிறழ்ச்சி யற்ற மதிப்பீடாகும். அதன் மாறுபாடானது

$$V(\hat{Y}_c) = \frac{1}{M} \sum_{1}^{M} (\bar{y}_i - \bar{Y})^2 = \sigma_b^2$$

என்பதாகும்.

'*c*' கீழ்குறி, (subscript), திரள் மாதிரி எடுப்பிணைக் குறித்து, σ_b^2 என்பது திரள்—இடைப்பட்ட மாறுபாடு (between clusters) ஆகும்.

ஆனால் ஒரே ஒரு திரளின் அடிப்படையில் மதிப்பீட்டின் மாதிரிசார் மாறுபாட்டு அளவினைத் தீர்மானிப்பது கடினம். எனவே, இரண்டுக்கு மேற்பட்ட அளவில் திரள்கள் தெரிவு செய்யப்படுவது சாலச் சிறந்தது.

மாதிரிமுறை சார்ந்த மாறுபாட்டின் அடிப்படையில், திரள் முறையானது. சமவாய்ப்பு முறையைவிடக் குறைந்த அளவில் திறன் உடையதாக (less efficient) அமையும் எனப் பின்பு நிறுவுவோம். சில குறிப்பிட்ட நிலையில் இரண்டு முறைகளும் சம அளவில் பயன்படும்.

(ii) '*m*' திரள்களைத் தேர்ந்தெடுத்தல்

ஒவ்வொன்றிலும் N_i $(i : 1, 2 \ldots m)$ உறுப்புகள் உள்ள *M* திரள்களினின்று, முதல் கட்டமாக *m* திரள்களைத் தேர்ந்தெடுக்கிறோம். இவை முதல்நிலை மாதிரிக் கூறுகள் (primary sampling units) எனப்படும். அடுத்த கட்டத்தில் *i*-ஆவது திரளிலிருந்து n_i $(i : 1, 2 \ldots m)$ இரண்டாம் நிலைக் கூறுகளை (secondary sampling units) தேர்ந்தெடுக்கவேண்டும். {இவ்வாறாக, இரு வேறு கட்டங்களில் அல்லது படிகளில் மாதிரித் தேர்வு நடைபெறுவதால், இதனை 'இரு கட்ட' அல்லது 'இருபடி' மாதிரிமுறை (two-stage sampling) என்றும் கூறுகிறோம்}.

இதில் n_i மதிப்பானது எல்லாத் திரள்களிலும் சமமாகவோ அல்லது வேறு வேறாகவோ அமைந்திடலாம். உதாரணமாக, ஒரு குறிப்பிட்ட பகுதியில் ஆய்வு நடத்தவேண்டியிருப்பின், குறிப்பாகக் கிராமங்களைக் கருதுகையில், கிராமங்களைத் திரள்களாகக் கொண்டு, அவற்றுள் *m* கிராமங்களைத் தேர்ந்தெடுத்து, பிறகு அக் கிராமங்களில் உள்ள குடும்பங்களிலிருந்து, அடுத்த கட்டத்தில் $n_1, n_2, n_3 \ldots$ என்ற அளவுகளில் மாதிரிகள் எடுக்கப்படலாம். நடைமுறைக்கு இது மிக எளிது.

மாறுபடும் திரள் அளவு (size of cluster) நிலையில், சம வாய்ப்பு முறை மாதிரி எடுத்தல் :

ஈடு செய்யப்படாத முறையின்மூலமாக, முழுமைத் தொகுதி கூட்டு மொத்தத்தினை (population total) மதிப்பிடல் :

முதல் கட்டத்தில், m திரள் மொத்தங்கள் மதிப்பிடப்பட வேண்டும். அடுத்த கட்டமாக, இத் திரள் கூட்டு மொத்தங்களைப் பயன்படுத்தி, M திரள்களின் மொத்தத்தை மதிப்பீடு செய்ய வேண்டும்.

$\hat{y}$ என்பது m மாதிரித் திரள்களைக் கொண்டு மதிப்பிடப் பெறும் கூட்டுமொத்தம் எனில்,

$$y = \sum_{i} Y_i$$

இது, முழுமைத் தொகுதித் திரள்களின் வெவ்வேறு மொத்த மதிப்பீடுகளின் ($\hat{Y}_i$, i : 1, 2...m) கூட்டுத்தொகையாகும். இதில்

$$\hat{Y}_i = N_i \left[\frac{1}{n_i} \sum_{j=1}^{n_i} y_{ij} \right] \text{ ஆகும்.}$$

N_i என்பது -iஆவது முழுமைத் தொகுதித் திரள் அளவு.

மேற்கண்ட வரையறையின் அடிப்படையில், முழுமைத் தொகுதியின் மொத்த அளவானது,

$$\hat{Y} = \frac{M}{m} \sum_{1}^{m} \hat{Y}_i$$

$$= \frac{M}{m} \cdot \sum_{i=1}^{m} \frac{N_i}{n_i} \sum_{j=1}^{n_i} y_{ij}$$

என்று மதிப்பீடு செய்யப்படும்.

[∵ $\bar{y}_i$ என்ற i-ஆவது மாதிரித் திரளின் சராசரியின் மூலமாக, Y_i மதிப்பிடப்பட்டு, பிறகு எல்லாத் திரள்களையும் இணைத்து, $\bar{y} = \frac{1}{m} \sum_{i} \bar{y}_i$ என்று பெற்று, இதிலிருந்து, $\hat{Y}$ க்கு ஒரு மதிப்பீடு பெறப்பட்டது.]

சுருக்கமாக,

$$\hat{Y} = M\left[\frac{1}{m}\cdot \sum_{i} N_i \left(\frac{1}{n_i}\sum_{j} y_{ij}\right)\right]$$

$$= M\left[\frac{1}{m}\sum_{i=1} N_i\,(\hat{\bar{y}}_i)\right].$$ இது முதல் கட்டம்.

$$\therefore\ \hat{Y} = M\left[\frac{1}{m}\sum_{i}\hat{Y}_i\right] = M.\,(\hat{\bar{Y}}).$$ [இது 2-ம் கட்டம்.]

வழிமுறை

முதலில். $\bar{y}_i$ மதிப்பினைக் கணக்கிட்டு, அதை N_i மூலமாகப் பெருக்கி, ($N_i\,\bar{y}_i$), $\hat{Y}_i$ காணப்படும்.

அடுத்து, முதல்நிலைத் திரள் சராசரியான $\hat{\bar{Y}}$ மதிப்பினைக் கணக்கிட்டு, $M.\,\hat{\bar{Y}}$-ன் மதிப்பைக் காணுவதன்மூலம், $\hat{Y}$ பெறப்படுகிறது.

எனவே, எளிய ராண்டம் முறையானது இரு கட்டங்களில் பயன்படுத்தப்படுகிறது.

தேற்றம் :

$$\hat{Y} = \frac{M}{m}\sum_{i}\frac{N_i}{n_i}\sum_{j} y_{ij}$$ என்ற மதிப்பீடானது, Y-ன் பிறழ்ச்சியற்ற மதிப்பீடாகும்.

அதாவது $E\,(\hat{Y}) = Y$ என நிருபிக்கவேண்டும்.

மாதிரிமுறையானது இரு கட்டங்களில் செயல்படுவதால் எதிர்பார்க்கும் மதிப்பினை (expected value) இரு வழிகளில் காண வேண்டும்.

$\hat{Y}$-ன் எதிர்பார்க்கும் மதிப்பினை,

$E\,(\hat{Y}) = E\ \ E_i\,(\hat{Y})$ என எழுதலாம்.

இதில் E_i என்பது முதல் கட்டத்தில் கணக்கிடப்படும் எதிர்பார்க்கும் மதிப்பாகவும், E_j என்பது 2-ம் கட்டத்தில் எதிர்பார்க்கும் மதிப்பாகவும் கொள்ளப்படும்.

அதாவது, E ஆனது M-முதல் நிலை மாதிரி உறுப்புகளைக் கொண்டு பெறப்படும் எதிர்பார்க்கும் மதிப்பு ; E_i ஆனது, N_i மாதிரி உறுப்புகளைக் (கூறுகளைக்) கொண்டு பெறப்படும் எதிர்பார்க்கும் மதிப்பு என்று அறியப்படும்.

எனவே, i-என்ற முதல்நிலைக் கூறு கொடுக்கப்பட்டிருப்பின், [i-நிலையாக],

$$E(\hat{Y}) = E_i\, E_j\, (\hat{Y})$$

$$= E_i\, E_j \left[\frac{M}{m} \sum_i \frac{N_i}{n_i} \sum_j y_{ij}\right]$$

$$= E_i \left[\frac{M}{m} \sum_i \frac{N_i}{n_i} \sum_j E_j(y_{ij})\right]$$

கொடுக்கப்பட்ட i-ஆவது முதல்நிலைக் கூறினில் (திரளில்) இருந்து n_i-அளவுள்ள ராண்டம் மாதிரி எடுக்கப்படுவதால்,

$E_j(y_{ij}) = \bar{Y}_i$ என்று கண்டு,

$$E(\hat{Y}) = E_i \left[\frac{M}{m} \sum_i \frac{N_i}{n_i} \sum_{j=1}^{n_i} \bar{Y}_i\right]$$

$$= E_i \left[\frac{M}{m} \Sigma \frac{N_i}{n_i} \cdot (n_i \cdot \bar{Y}_i)\right]$$

$$\therefore \; E(\hat{Y}) = E_i \left[\frac{M}{m} \sum_i N_i \bar{Y}_i\right] = E\left[\frac{M}{m} \sum_{i=1}^{m} y_i\right]$$

y_i என்பது i ஆவது முதல் நிலைக் கூறின் மொத்தமாகும். அதாவது, எடுக்கப்பட்ட i-ம் ராண்டம் மாதிரியின் கூட்டு மொத்தம்.

இனி, M முதல் நிலைக் கூறுகள் (திரள்கள்) வாயிலாக எதிர்பார்க்கும் மதிப்பினைக் கணக்கிடவேண்டியிருப்பதால் y_i-யினை ஒரு வாய்ப்பு மாறி எனக் கருதி (random variable), $i = 1, 2 \ldots M$ என்ற படியாக, ஒவ்வொரு திரளும் தெரிவுசெய்யப்பட $\frac{1}{M}$ என்ற சமவாய்ப்பு நிலையில்,

$$E(\hat{Y}) = E_i\left[\frac{M}{m}\sum_{i} y_i\right]$$

$$= \frac{M}{m} \sum_{i} E_i(y_i)$$

$$= \frac{M}{m} \cdot \sum_{i}^{m} \sum^{M} \frac{1}{M} \cdot y_i$$

$$= \frac{M}{m} \cdot m \cdot \frac{1}{M} \sum_{1}^{M} y_i = Y \quad \text{ஆகிறது.}$$

எனவே,

$E(\hat{Y}) = Y$ என்ற அளவில், $\hat{Y}$ ஆனது Y-ன் பிறழ்ச்சியற்ற மதிப்பீடாகிறது.

குறியீடு :

Y_i = i ஆவது திரளின் மொத்தம்.

$\bar{Y} = \frac{Y}{M}$ = திரள்களின் கூட்டுச் சராசரி.

$\bar{Y}_i = \frac{Y_i}{N_i}$ = i-ஆவது திரளின் சராசரி மதிப்பீடு.

(iii) சராசரி

இதன் அடிப்படையில், முழுமைத் தொகுதிச் சராசரியான $\bar{Y}$-ஐ பின்வருமாறு மதிப்பீடு செய்யலாம். அதாவது, $\bar{Y}$-ன் மதிப்பீட்டைப் பெறுதற்கு (estimate), மொத்தமான Y-ன் மதிப்பீட்டினை, N-ஆல் வகுக்க வேண்டும்.

$$\therefore \hat{\bar{Y}} = \frac{\hat{Y}}{N} = \frac{1}{N} \cdot \frac{M}{m} \sum_{i} \frac{N_i}{n_i} \cdot \sum_{j} y_{ij}$$

[கவனிக்க : $\hat{\bar{Y}}$ மதிப்பீடானது, மாதிரிச் சராசரி மதிப்பு அல்ல.]

ஆனால், இரு கட்ட மாதிரிச் சராசரி $\bar{y} = \frac{y}{n}$ என்பதே ஆகும்.

எனவே, திரள் முறை மாதிரித் தேர்வில், Y-மதிப்பீட்டிலிருந்து, $\bar{\bar{Y}}$-ன் மதிப்பீடானது பெறப்படுகிறதென அறியலாம்.

முன்புபோலவே, $\hat{\bar{\bar{Y}}}$ ஆனது $\bar{\bar{Y}}$-ன் பிறழ்ச்சியற்ற மதிப்பாக அமையும். ஏனெனில், $\hat{Y}$ ஆனது Y-ன் பிறழ்ச்சியற்ற மதிப்பீடாக நிறுவப்பட்டது.

$$\therefore \quad E(\hat{\bar{\bar{Y}}}) = \frac{E(\hat{Y})}{N} = \frac{1}{N}(Y) = \bar{\bar{Y}}.$$

என்பது முடிவாகும்.

(iv) $\hat{Y}$-ன் மாறுபாட்டு அளவு மற்றும் அதன் மதிப்பீடு :
(Sampling variance of $\hat{Y}$ and its Estimate).

முதலில், $V(\hat{Y}_c)$-யினைக் கருதுவோம்.

c என்பது திரள் மாதிரி முறையினைக் குறிப்பிடும்.

$\hat{Y}$ மதிப்பீட்டுத் திட்பத்தினை (precision) அறிதற்கு, அதன் மாதிரி சார் மாறுபாட்டை (sampling variance) காணவேண்டும். $V(\hat{Y}_c)$ காணப்படுகையில், இரு கட்டங்களைக் கருதவேண்டும். மாறுபாடானது, இரு காரணங்களால் (source) ஏற்படக் காணலாம்.

(1) திரள்களின் இடையிட்ட மாறுபாடு (between clusters).

(2) ஒவ்வொரு திரளினின்றும் 2-ம் கட்டத் தேர்வுகள் நடைபெறுவதால் ஏற்படக்கூடிய திரள்களின் உள்ளீட்டு மாறுபாடு (within clusters).

∴ மொத்த அளவு மதிப்பீட்டின் முழு மாறுபாடானது

$$V(\hat{Y}_c) = M^2 \cdot \frac{M-m}{M} \cdot \frac{S_b^2}{m} + \frac{M}{m} \sum_i N_i^2 \cdot \frac{(N_i - n_i)}{N_i} \cdot \frac{S_i^2}{n_i}$$

என்று பெறப்படும்.

இதில்,

$$S_b^2 = \frac{1}{(M-1)} \sum_{i=1}^{M} (Y_i - \bar{Y})^2 \text{ (இடையிட்ட மாறுபாடு)}$$

மேலும்,

$$S_i^2 = \frac{1}{(N_i - 1)} \sum_{j=1}^{N_i} (y_{ij} - \bar{\bar{Y}}_i)^2 \text{ (உள்ளீட்டு மாறுபாடு).}$$

Y_i என்பது i-என்ற திரளின் மொத்த அளவு.

$\bar{\bar{Y}}_i = \frac{Y_i}{N_i}$ என்பது i-என்ற திரளினது சராசரி.

∴ S_b^2 என்பது திரன் மொத்தங்களின் (totals) இடையீட்டு மாறுபாடாகும். ($\bar{Y}$-லிருந்து Y_i-மதிப்புகளின் மாறுபாடு).

இதேபோல், S_i^2 என்பது, i-என்ற திரளின் உள்ளிட்ட மாறுபாட்டு அளவாக அறியப்படும்.

($\bar{\bar{Y}}_i$-லிருந்து y_{ij} மதிப்புகளின் மாறுபாடு)

இதைப் பெறும் வழிமுறை :

வரையறையின்படி,

$$V(\hat{Y}_c) = E[\hat{Y} - E(\hat{Y})]^2 = E[\hat{Y} - Y]^2$$

$(\hat{Y} - Y)^2$ மதிப்பை முதலில் பெறவேண்டும். மாதிரித் தேர்வானது, இரு கட்டங்களில் கருதப்படுவதால், முதல் கட்டத்தை நிலையாக எடுத்துக்கொண்டு (fixed), இரண்டாவது கட்டத்தில் துவங்கி, ஒரு நிலை முடிவினைப் பெற்று, அதன்பின் முதலாவது கட்டத்தில் மேற்கொள்ளும் வழிமுறைமூலம் இறுதி முடிவைப் பெறுகிறோம்.

$$(\hat{Y} - Y)^2 = \left[\frac{M}{m} \sum_{i} \hat{Y}_i - Y\right]^2$$

$$= \left[\left(\frac{M}{m} \sum_i \hat{Y}_i - \frac{M}{m} \sum_i Y_i\right) + \left(\frac{M}{m} \sum_i Y_i - Y\right)\right]^2$$

$$= \left(\frac{M}{m} \sum_i Y_i - Y\right)^2 + 2\left(\frac{M}{m} \sum_i Y_i - Y\right) \left(\frac{M}{m} \Sigma \hat{Y}_i - \frac{M}{m} \Sigma \hat{Y}_i\right) + \left(\frac{M}{m} \Sigma \hat{Y}_i - \frac{M}{m} \Sigma Y_i\right)^2$$

$$= \left(\frac{M}{m} \Sigma Y_i - Y\right)^2 + 2\left(\frac{M}{m}\right) \left(\frac{M}{m} \sum_i Y_i - Y\right) \left(\sum_i (\hat{Y} - Y_i\right) + \left(\frac{M}{m}\right)^2 \left[\sum_i (\hat{Y} - Y_i)\right]^2$$

$$= \left(\frac{M}{m} \Sigma Y_i - Y\right)^2 + 2\left(\frac{M}{m}\right) \left(\frac{M}{m} \Sigma Y_i - Y\right) \left(\Sigma (\hat{Y} - Y_i)\right) + \left(\frac{M}{m}\right)^2 \Sigma (\hat{Y}_i - Y_i)^2 + \left(\frac{M}{m}\right)^2 \sum_{\neq j} (\hat{Y}_i - Y_i) (\hat{Y}_j - Y_j)$$

அதாவது

$$(\hat{Y} - Y)^2 = (1) + (2) + (3) + (4)$$

என எழுதுவோம்.

இனி, i என்ற முதல் நிலைக் கூறினை (திரள்) நிலையாகக் கொண்டு, எதிர்பார்க்கும் மதிப்பு (E) காண்போம் :

அதாவது, $$E_j(\hat{Y} - Y)^2 = E_j(1) + E_j(2) + E_j(3) + E_j(4)$$

என்றாகும். வலப்புற உறுப்புகளைத் தனித்தனியாக ஆராய்வோம்:

$$\therefore \quad E_j\ (1) = E_j \left[\frac{M}{m} \sum_i (Y_i - Y)\right]^2$$

$$= \left[\frac{M}{m} \sum_i Y_i - Y\right]^2$$

ஏனெனில் இதில், இரண்டாம் கட்ட உறுப்புகள் வருவதில்லை.

$$E_j(2) = E_j\left[2\left(\frac{M}{m}\right)\left(\frac{M}{m}\sum_i \overline{\overline{Y}}_i - \overline{\overline{Y}}\right)\sum_i (\hat{Y}_i - Y_i)\right]$$

$$= 2\left(\frac{M}{m}\right)\left[\frac{M}{m}\sum_i (Y_i - Y)\right] \cdot E_j \sum_i (\hat{Y}_i - Y_i)$$

($\because$ $\hat{Y}_i$ என்ற மாதிரி மதிப்பீட்டு அளவை (statistic) ஆனது, இரண்டாம் நிலைக் கூறுகள் தெரிவின் அடிப்படையில் கணிக்கப் படுகிறது. அதாவது, i-ஆம் திரளின் 'மொத்த' மதிப்பீடானது, n_i அளவுள்ள சமவாய்ப்பு மாதிரி அடிப்படையில் அமைகிறது.)

ஆனால், $E(\hat{Y}_i) = Y_i$ ஆகும். (சமவாய்ப்பு மாதிரியில்).

$\therefore$ மேலும் $E(\Sigma X) = \Sigma E(x)$ என்ற நியதி அடிப் படையில்,

$$E_j \sum_i (\hat{Y}_i - Y_i) = \sum_j E_i(\hat{Y}_i - Y_i) = 0$$ என

அமைய,

$E_j(2) = 0$ ஆகிறது.

இதேபோன்று E (4)-ம் அமையக் காணலாம்.

அதாவது $E_j(4) = 0$.

இறுதியாக,

$$E_j(3) = E_j\left[\left(\frac{M}{m}\right)^2 \cdot \Sigma(\hat{Y}_i - Y_i)^2\right]$$

$$= \left(\frac{M}{m}\right)^2 \sum_i E_j(\hat{Y}_i - Y_i)^2$$

i-நிலையானதாக இருக்க, திரள்களினின்று எடுக்கப்படும் துணைக் கூறுகள், எளிய சமவாய்ப்பு மாதிரியை அமைப்பதாகக் கொண்டு,

$$E_j(\hat{Y}_i - Y_i)^2 = N^2 \cdot \frac{N_i - n_i}{N_i} \cdot \frac{S_i^2}{n_i}$$

$$\left[S_i^2 = \frac{1}{N_i - 1} \sum_{j=1}^{N_i} (Y_{ij} - \bar{\bar{Y}}_i)^2\right]$$

எனப் பெறலாம்.

எனவே, தேவையான மாறுபாடு

$$E_j(\bar{Y} - Y)^2 = \left[\frac{M}{m} \sum_i Y_i - Y\right]^2 + \left(\frac{M}{m}\right)^2 \sum_i N_i^2 \cdot \frac{N_i - n_i}{N_i} \cdot \frac{S_i^2}{n_i}$$

என்று பெறப்படும்.

அடுத்த நிலையில், M திரள்களினின்று எடுக்கப்படும் எல்லா மாதிரித் திரள்களையும் கருத்திற்கொண்டு, எதிர்பார்க்கும் மதிப்பானது கணிக்கப்படவேண்டும்.

எனவே, $E_i(\hat{Y} - Y)^2$-ல் உள்ள i ஆனது மாறுகிறது.

மேலும் $E_j(\hat{Y}-Y)^2$ என்பது வாய்ப்பு மாறியாகிறது.

$$\therefore \quad E_i[E_j(\hat{Y}-Y)^2] = E_i\left[\frac{M}{m}\sum_i Y_i - Y\right]^2$$

$$+ E_i\left(\frac{M}{m}\right)^2 \sum_i N_i^2 \cdot \frac{N_i-n_i}{N_i} \cdot \frac{S_i^2}{n_i}.$$

[E_j (1) & E_j (3) ஆகியவை பிரதியிடப்பட்டன.]

வலப்புற இரண்டாம் உறுப்பாவது,

$$E_i\left(\frac{M}{m}\right)^2 \sum_i N_i^2 \cdot \frac{N_i-n_i}{N_i} \cdot \frac{S_i^2}{n_i}$$

$$= \left(\frac{M}{m}\right)^2 \cdot \sum_i E\left[N_i^2 \cdot \frac{N_i-n_i}{N_i} \cdot \frac{S_i^2}{n_i}\right]$$

$$= \left(\frac{M}{m}\right)^2 \sum_i \left\{\frac{1}{M}\sum_{i=1}^{M}\left[N_i^2 \cdot \frac{N_i-n_i}{N_i} \cdot \frac{S_i^2}{n_i}\right]\right\}$$

($\because$ மொத்தம் M மதிப்புகள் காணப்பட்டு, ஒவ்வொன்றுக்கு $\frac{1}{M}$ என்ற அளவில் சமவாய்ப்பு அமைகிறது.)

இது, i என்ற திரளின் உள்ளீட்டு மாறுபாட்டை அளிக்கும்.

முதல் உறுப்பான

$$E_i\left[\frac{M}{m}\sum_i Y_i - Y\right]^2 \text{ என்பதில், } \frac{M}{m}\sum_i Y_i$$

என்ற பகுதியை, m திரள்களின் அடிப்படையில் பெறப்பட்ட Y-ன் மதிப்பீடாகக் கொள்ளலாம். எனவே, மொத்தங்களின் மதிப்பீடு பற்றிய சமவாய்ப்பு மாதிரி முடிவுகளினைப் பயன்படுத்தி,

$$E_i\left[\frac{M}{m}\sum_i Y_i - Y\right]^2 = M^2 \cdot E_i\left[\frac{1}{m}\Sigma Y_i - \overline{Y}\right]^2$$

$$= M^2 \cdot \frac{M-m}{M} \cdot \frac{1}{m}\,\frac{\Sigma\,(Y_i - \overline{Y})^2}{M-1}$$

என்று பெறப்படும்.

இறுதியாக,

$$V\,(Y_c) = \frac{M^2}{m} \cdot \frac{M-m}{M} \cdot S_b^2 + \frac{M}{m}\sum_1^M N_i^2 \cdot \frac{N_i - n_i}{N_i} \cdot \frac{S_i^2}{n_i}.$$

(S_b^2, S_i^2 முன்பே கொடுக்கப்பட்டவை).

என்று நிறுவுகிறோம்.

துணை முடிவு 1

சம வாய்ப்பு முறைப்படி, $\frac{S_i^2}{n_i}$ என்பது $\overline{Y}_i$-யினைப் பற்றிய, $\bar{y}_i$-ன் மாதிரிமுறைசார் மாறுபாடாகும். $\bar{y}_i$ ஆனது, N_i யிலிருந்து எடுக்கப்பட்ட n_i அளவுள்ள மாதிரியினின்று பெறப்பட்ட சராசரியாகும். மேலும், $\frac{N_iS_i^2}{n_i}$ என்பது, $\hat{Y}_i = N_i\ \bar{y}_i$-ன் மாதிரிசார் மாறுபாடாகும். இதிலிருந்து, $N_i = n_i$ என்று அமைந்திடின், $\bar{y}_i = \overline{Y}_i$ எனவும், $\hat{Y}_i = Y_i$ எனவும் வருவதால், Y_i-ன், Y_i பற்றிய மாதிரிமுறைசார் மாறுபாடானது பூச்சியமாகிறது.

அதாவது, முதல்நிலைக் கூறுகளில் (அதாவது திரள்களில்) உள்ள எல்லா மாதிரிக் கூறுகளும் தெரிவு செய்யப்படின்

($N_i = n_i$ என்றபடி), $\Sigma N_i^2\ \frac{(N_i - n_i)}{N_i}.\ \frac{S_i^2}{n_i} = 0$

என்று அமைந்திடும்.

எனவே,

$$V\,(\hat{Y}_c) = M^2\ \frac{M-m}{M}.\ \frac{S_b^2}{m}.$$

($\because$ முதல்நிலைக் கூறுகளின் உள்ளீட்டு மாறுபாடு பூச்சிய மாகும்.) இந்நிலையில், திரள்களின் இடைப்பட்ட (between) மாறுபாடு மட்டுமே இருக்கக் கூடும்.

துணை முடிவு 2

இதற்கு மாறாக, இரண்டாம் நிலைக்கூறுகள், எல்லா M முதல் நிலைக் கூறுகள் அல்லது திரள்களினின்றும் எடுக்கப்படின் ($M = m$ எனும்படி), திரள்களின் இடைப்பட்ட மாறுபாடானது பூச்சியமாக, $V(\hat{y})$ ஆனது, உள்ளீட்டு மாறுபாட்டினை மட்டும் கொண்டிருக்கும்.

அதாவது,

$$V(\hat{Y}_c) = \frac{M}{m} \sum_{i=1}^{M} N_i^2 . \frac{N_i - n_i}{N_i} . \frac{S_i^2}{n_i}$$

$$= \sum_{i=1}^{M} N_i^2 . \frac{N_i - n_i}{N_i} . \frac{S_i^2}{n_i} .$$

இது, படுகைமாதிரி முறையில் பெறப்பட்ட $V(\hat{Y}_{st})$ - யின் மதிப் பாகும் என்பது குறிப்பிடத்தக்கது.

(v) $V(\hat{Y}_c)$ யின் மதிப்பீட்டு அளவை

(Estimation of Variance of $\hat{Y}_c$)

நடைமுறையில், ஒரு முழுமைத் தொகுதியானது பெரிய அளவினதாக அமைவதால், $V(\hat{Y}_c)$ யினை நேரடியாக மதிப்பிடுவது இயலாது. எனவே, அதற்கு ஒரு மதிப்பீடு காண வேண்டியுள்ளது. $V(\hat{Y}_c)$ யின் மதிப்பீட்டு அளவையினைத் தீர்மானிக்கையில், அதன் இருவேறுப் பகுதிகளைப் பற்றி அறிய வேண்டும். S_b^2, S_i^2 என்ற இடையிட்ட மாறுபாடு, உள்ளீட்டு மாறுபாடு ஆகியவைபற்றி நன்கு அறிதல் வேண்டும். $\therefore$ S_b^2, S_i^2 இவற்றின் மாதிரி மதிப்பீடுகளைப் பெற்று, தக்கபடி பிரதியிட,

$$\hat{V}(\hat{Y}_c) = M^2 . \frac{M-m}{M} . \frac{s_b^2}{m} + \frac{M}{m} \sum_{i=1}^{m} N_i^2 . \frac{N_i - n_i}{N_i} . \frac{s_i^2}{n_i}$$

எனப் பெறலாம்.

இதில்

$$s_b^2 = \frac{1}{m-1} \sum_{i=1}^{m} (\hat{Y}_i - \hat{\bar{Y}})^2$$

$$s_i^2 = \frac{1}{n_i - 1} \sum_{j=1}^{n_i} (y_{ij} - \bar{\bar{y}}_i)^2$$

குறிப்பு : $\hat{V}(\hat{\bar{y}}_c)$ யில், வலப்புற இரண்டாம் உறுப்பில் வரும் Σ ஆனது m திரள்களைப் பற்றி அமைவது கவனிக்கத்தக்கது (M அல்ல).

(i) s_b^2 ஆனது m-திரள்களினிடைப்பட்ட $\hat{Y}_i$-ன் மாறுபாட்டினைக் காட்டுவதாகும்.

$$\hat{Y}_i = N_i \cdot \bar{\bar{y}}_i \left(\because \ \bar{y}_i = \frac{y_i}{n_i} \right)$$

என்பது, i-ஆவது திரளின் மொத்தத்திற்கு உள்ள மதிப்பீட்டு அளவை.

$$\hat{\bar{Y}} = \frac{1}{m} \sum_{1}^{m} \hat{Y}_i \quad \text{என்பது} \quad \hat{Y}_i \quad (i : 1, 2 \ldots m)$$

மதிப்புகளின் சராசரி.

(ii) s_i^2 என்பது, i ஆவது திரளிலிருந்து எடுக்கப்படும் மாதிரியின், y_{ij} மாறியினது, உள்ளீட்டு மாறுபாடாகும். N_i அளவிலிருந்து எடுக்கப்படும் n_i-அளவுள்ள சமவாய்ப்பு மாதிரியின் சராசரி $\bar{y}_i$ ஆவதால், நியதிப்படி, s^2 என்பது S^2-ன் பிறழ்ச்சியற்ற மதிப்பீடாவது காணப்படும்.

$\therefore \quad E(s_i^2) = S_i^2.$

ஆனால் s_b^2 என்பது S_b^2-ன் பிறழ்ச்சியற்ற மதிப்பீடு அல்ல. ஏனெனில்,

$$E(s_b^2) = S_b^2 + \frac{1}{M} \cdot \sum_{1}^{M} N_i^2 \cdot \frac{N_i - n_i}{N_i} \cdot \frac{S_i^2}{n_i}$$

என்று அமையும்.

$\hat{\bar{\bar{Y}}}_c$-ன் மாதிரிசார்ந்த மாறுபாடும், அதன் மதிப்பீடும், இம் முடிவுகளைப் பயன்படுத்தி, பின்வருமாறு பெறப்படுகின்றன :

$$V(\hat{\bar{\bar{Y}}}_c) = V\left(\frac{\hat{Y}_c}{N}\right) = \frac{1}{N^2} \cdot V(\hat{Y}_c)$$

$$\therefore \quad V(\hat{\bar{\bar{Y}}}_c) = \frac{1}{N^2}\left[M^2 \cdot \frac{M-m}{m} \cdot \frac{S_b^2}{m} + \frac{M}{m} \cdot \sum_1^M N_i^2 \cdot \frac{N_i - n_i}{N_i} \cdot \frac{S_i^2}{n_i}\right]$$

இதன் மாதிரி மதிப்பீடாவது,

$$\hat{V}(\hat{\bar{\bar{Y}}}_c) = \frac{1}{N^2}\left[M^2 \cdot \frac{M-m}{m} \cdot \frac{S_b^2}{m} + \frac{M}{m}\sum_1^m N_i^2 \cdot \frac{N_i - n_i}{N_i} \cdot \frac{s_i^2}{n_i}\right]$$

குறிப்பு : சில நிலைகளில், $\hat{\bar{\bar{Y}}}$, $\hat{V}(\hat{\bar{\bar{Y}}}_c)$ ஆகியவற்றை மதிப்பிட, N-ன் மதிப்புத் தெரியாமல் இருக்கக்கூடும். அதற்காக, (i) எல்லா முதல்நிலைக் கூறுகளையும் (திரள்களை) சம அளவினதாகக் கருதி, இவற்றிலிருந்து மாதிரிகள் எடுத்து, அவற்றினின்றும் தெரிவு செய்யப்படும் இரண்டாம் நிலைக் கூறுகளின் அளவினையும் சமமாக எடுத்துக்கொள்ளலாம். அல்லது, (ii) ஒரு தோராயமான அளவில் N மதிப்பைப் பெறலாம். இவை இரண்டினையும் பின்பு காண்போம்.

மூன்றாவதாக வரும் முறையானது ஒரு முக்கியமான மாதிரி முறையைத் தருவதால், அதை விரிவாக வேறொரு பகுதியில் ஆராய்வோம். அதன்படி, முதல்நிலைக் கூறுகளை அல்லது திரள்களை அளவிற்கேற்ற வாய்ப்பு விகிதத்தில் தெரிவு செய்தல் வேண்டும்.

இனி, திரள்மாதிரி முறையின் ஒரு முக்கிய நிலையினை சற்றுக் காண்போம் :

(vi) திரள் அளவுகளும், துணைக்கூறு அளவுகளும் சமமான நிலை :
(Cluster sizes & Sub-sample sizes equal)

சில ஆய்வுப் பிரச்னைகளில், கருதப்படும் திரள் அளவுகள் N_i சமம் எனவும், இத் திரள்களினின்று தெரிவு செய்யப்படும் துணைக் கூறு அளவுகளும் (n_i) சமம் எனவும் கொள்ளலாம். துணைக் கூறுகள், இறுதித் தொகுப்புகள் அல்லது திரள்கள் என்று கருதப்படும். இந்நிலைப்படி,

$$N_i = \frac{N}{M} = N' \text{ என்றும்}$$

$$n_i = \frac{n}{m} = n' \text{ என்றும்}$$

ஆகும்.

இந்த முறையானது, விவசாயத் துறைகளில் மற்றும் வனத் துறைகளில் மேற்கொள்ளப்படும் ஆய்விற்கு மிக உதவும்.

உதாரணமாக, ஒரு பயிரின் விளைவினைப்பற்றி அறிய, மொத்தப் பரப்பானது, 5 ஏகர் அளவுள்ள சிறு பரப்புகளாகப் பிரிக்கப்பட்டு மொத்தம் M-பரப்புக்கள் இருப்பின், ஒவ்வொரு 5-ஏகர் பரப்பும் சம அளவில் 20 பாத்திகளாக (plots) பிரிக்கப் பட்டு, $(N' = 20)$ இரு கட்டத் தேர்வு நடைபெறுவதாகக் கொள்வோம். முதலில் m பரப்புகள் தெரிவு செய்யப்படும். ஒவ்வொரு பரப்பினின்றும், அடுத்த கட்டத்தில், $n' = 5$ என்ற அளவில் ஒரே சீரான பாத்திகளைத் தெரிவு செய்ய, இந்நிலையில், பாகுபாடு எளிதாகக் காணலாம்.

முன்புபோலவே,

$$\hat{Y} = \frac{M}{Nm} \sum_{1}^{m} \frac{N'}{n'} \cdot \sum_{j:1}^{n'} y_{ij}$$

$$= \frac{1}{mn'} \sum_{m} \sum_{n'} y_{ij}$$

என்பது, மாதிரிச் சராசரி ஆகும்.

அதாவது, $m.n' = n$ (மாதிரி அளவு).

மேலும் $E(\hat{\bar{Y}}) = \bar{Y}$ என்று காணலாம்.

($\bar{Y}$-ஐ மதிப்பீடு செய்ய, N தெரிந்திருக்கவேண்டிய அவசியமில்லை.)

$\hat{\bar{Y}}_c$-ன் மாறுபாடு :

$$V(\hat{\bar{Y}}_c) = \frac{1}{N^2}\left[M^2 \cdot \frac{M-m}{M} \cdot \frac{S_b^2}{m} + \frac{M}{m} \cdot \sum_M N_i^2 \cdot \frac{N_i - n_i}{N_i} \cdot \frac{S_i^2}{n_i}\right]$$

என்பதன்மூலம், தக்கபடி பிரதியிடல் வழியாகப் பெறலாம்.

$$\frac{1}{N^2} \cdot M^2 \cdot \frac{M-m}{M} \cdot \frac{S_b^2}{m} = \frac{M-m}{M} \cdot \frac{1}{m} \cdot \frac{S_b}{(N')^2}$$

மேலும், $S_b^2 = \frac{1}{M-1} \sum_M (Y_i - \bar{Y})^2$

$$\therefore \quad \frac{S_b^2}{(N')^2} = \frac{1}{(M-1)} \sum_M \left(\frac{Y_i}{N'} - \frac{\bar{Y}}{N'}\right)^2$$

$$= \frac{1}{(M-1)} \cdot \sum_M (\bar{Y}_i - \bar{Y})^2$$

இது, முதல் நிலைக் கூறுகளின் சராசரிகளின் (திரள்களின்) இடைப்பட்ட மாறுபாட்டினை அளிக்கும்.

$\bar{Y}_i$ என்பது, i-ஆம் திரளில் உள்ள சராசரியாகவும், $\bar{Y}$ என்பது, மொத்தத்தில் அமையும் சராசரியாகவும் ஆகின்றன.

$$S_{1b}^2 = \frac{S_b^2}{M-1} \Sigma (\bar{Y}_i - \bar{Y})^2 = \frac{S_b^2}{(N')^2} \qquad \text{...... (A)}$$

என்று எழுதுவோம்.

வலப்புற இரண்டாம் உறுப்பானது $[V(\hat{\bar{Y}})$-ல்$]$,

$$\frac{1}{N^2}\cdot\frac{M}{m}\sum_{M} N_i^2.\ \frac{N_i - n_i}{N_i}.\ \frac{S_i^2}{n_i}$$

$$= \frac{1}{N^2}\cdot\frac{M}{m}\cdot\frac{N^2}{M^2}\cdot\frac{N' - n'}{N'}.\ \sum_{M}\frac{S_i^2}{n_i}$$

$$= \frac{N' - n'}{N'}.\ \frac{1}{m}.\ \sum_{M}\frac{S_i^2}{mn'} \qquad \ldots\ldots \text{(B)}$$

என்று ஆகும்.

இதில், $\frac{1}{M}\sum_{M} S_i^2 = \frac{1}{M}\ \frac{1}{(N'-1)}\sum_{M}\sum_{N'}(y_{ij}-\bar{\bar{Y}})^2$

என்பது திரள்களின் உள்ளிட்ட இணைந்த மாறுபாடாகும்.

இதனை $S_{2i}^2 = \frac{1}{M}\Sigma S_i^2$ என்று குறிக்கலாம்.

எனவே, (A), (B) ஆகியவற்றைப் பிரதியிட்டு,

$$V(\hat{\bar{Y}}_c) = \frac{M-m}{M}.\ \frac{S_{1b}^2}{m} + \frac{N'-n'}{N'}.\ \frac{S_{2i}^2}{mn'}$$

என்று பெறலாம்.

அல்லது,

$$V(\hat{\bar{Y}}) = \frac{M-m}{M}.\ \frac{S_{1b}^2}{m} + \frac{N'-n'}{N'}.\ \frac{1}{M}\sum_{M}\frac{S_i^2}{mn'}$$

இதன் பிறழ்ச்சியற்ற மதிப்பீடாவது :

$$\hat{V}(\hat{\bar{Y}}) = \frac{M-m}{M}.\ \frac{s_{1b}^2}{m} + \frac{N'-n'}{N'}.\ \frac{1}{M}\sum\frac{s_i^2}{mn'}$$

இதில்,

$$s_{1b}^2 = \frac{1}{m-1}\sum_{m}(\hat{\bar{y}}_j - \bar{y})^2,\ \left(\bar{y}_i = \frac{y_i}{n'}\ ;\quad \bar{y} = \frac{y}{n}\right)$$

எனவும்

$$s_i^2 = \frac{1}{n'-1} \sum_{n'} (y_{ij} - \bar{y}_i)^2$$

எனவும் அமைந்திருப்பது குறிப்பிடத்தக்கது.

அல்லது,

$$\hat{V}(\hat{Y}_c) = \frac{M-m}{M}. \frac{s_{1b}^2}{m} + \frac{N'-n'}{N'}. \frac{1}{M} \sum_{m} \frac{s_i^2}{mn'}$$

$$= \frac{M-m}{M}. \frac{s_{1b}^2}{m} + \frac{N'-n'}{N'}. \frac{m}{M}. \frac{s_{2i}^2}{mn'}$$

என்றும் எழுதலாம்.

$$\left[s_{2i}^2 = \frac{1}{m} \sum_{m} s_i^2 = \frac{1}{m(n'-1)}. \sum_{m} \sum_{n} (y_{ij} - \bar{y}_i)^2 \right]$$

s_{2i}^2 = உள்ளீட்டு மாறுபாடுகளின் இணைப்பு மாறுபாடாகும்.

s_{1b}^2, s_{2i}^2 இவற்றை நடைமுறையில் கணக்கிட,

பின்வருமாறு கருதலாம் :

$$s_{1b}^2 = \frac{1}{m-1} \sum_{m} (\bar{y}_i - \bar{y})^2 = \frac{1}{m-1} (\Sigma \bar{y}_i^2 - m\bar{y}^2)$$

ஏனெனில், $\Sigma \bar{y}_i^2 = \left(\frac{1}{n'}\right)^2. \sum_{m} \left[\sum_{n'} y_{ij} \right]^2$

என்றும், $\bar{y}^2 = \left(\frac{1}{mn'}\right)^2. \left(\sum_{m} \sum_{n'} y_{ij} \right)^2$

என்றும் அமைந்திருக்கும்.

மேலும்,

$$s_{2i}^2 = \frac{1}{m(n'-1)}. \sum_{m} \sum_{n'} (y_{ij} - \bar{y}_i)^2$$

$$= \frac{1}{m(n'-1)} \sum_{m} \left[\sum_{n'} y_{ij}^2 - n' \bar{y}_i^2 \right]$$

$$= \frac{1}{m(n'-1)} \left[\sum_{m} \sum_{n'} y_{ij}^2 - n' \sum_{m} \bar{y}_i^2 \right]$$

$$\therefore \quad s_{2i}^2 =$$

$$\frac{1}{m(n'-1)} \left[\sum_{m} \sum_{n'} y_{ij}^2 - \frac{1}{n'} \sum_{m} \left(\sum_{n'} y_{ij} \right)^2 \right]$$

இத்தகைய நிலைகளிலும், முழுமைத் தொகுதி அளவான N-ன் மதிப்பு அறியப் பெற்றிருத்தல் அவசியமாகிறது.

எடுத்துக்காட்டு 1 :

பின்வரும் தொகுதியானது கொடுக்கப்படுகிறது.

திரள் எண்	y_{ij}
A	2, 3, 5
B	3, 5, 7
C	5, 7. 12

$m = 2$ என்ற அளவில்த் திரள்களை தெரிவுசெய்து. ஒவ்வொரு திரளிலும் $n_1=2$, $n_2=2$ என்ற அளவில் மாதிரிகளை எடுப்போம்.

எடுக்கக்கூடிய மாதிரிகள் :

திரள்களுள், மொத்தமாக $\binom{M}{m} = \binom{3}{2} = 3$ கூறுகள் எடுக்கலாம். இவற்றிலிருந்து, 2-ம் நிலைக் கூறுகளை, $\binom{3}{2}\binom{3}{2} = 9$ என்ற அளவில் எடுக்க, மொத்த மாதிரிகளின் எண்ணிக்கை $\binom{M}{m}\binom{3}{2}\binom{3}{2} = 27$ ஆகும்,

அவையாவன :

A	B	$\hat{Y}$	A	C	$\hat{Y}$	B	C	$\hat{Y}$
2, 3	3, 5 3, 7 5, 7	28·25 33·75 38·25	2, 3	5, 7 7, 12 5, 12	38·25 54·00 49·50	3, 5	5, 7 7, 12 5, 12	45·00 60·75 56·25
3, 5	3, 5 3, 7 5, 7	36·00 40·50 45·00	3, 5	5, 7 7, 12 5, 12	45·00 60·75 56 25	3, 7	5, 7 7, 12 5, 12	49 50 65·25 60·25
2, 5	3, 5 3, 7 5, 7	33·75 38·25 42·75	2, 5	5, 7 7. 12 5, 12	42·75 58·50 54·00	5, 7	5, 7 7, 12 5, 12	54·00 69·75 65·25
		336·50			459·00			526·00

(1) Y-ன் மதிப்பீடு :

முதல் மாதிரியான (2, 3, 3, 5) என்பதிலிருந்து,

$$\hat{Y} = \frac{M}{m} \sum_{1}^{m} \frac{N_i}{n_i} \sum_{j} y_{ij}$$

$$= \frac{3}{2} \left[\frac{N_1}{n_1} \sum_{j} y_{1j} + \frac{N_2}{n_2} \sum_{j} y_{2j} \right]$$

$$= \frac{3}{2} \left[\frac{N_1}{n_1} (y_{11} + y_{12}) + \frac{N_2}{n_2} (y_{11} + y_{22}) \right]$$

$$= \frac{3}{2} \left[\frac{3}{2} (2 + 3 + 3 + 5) \right]$$

$$= \frac{9}{4} \cdot 13 = 29{\cdot}25$$

இதேபோன்று, 27 மாதிரிகளிலும் $\hat{Y}$ மதிப்புகளைக் காணலாம். இவ்வாறு காணப்பட்டவற்றிலிருந்து $E(\hat{Y})$ மதிப்பிடலாம். வரையறைப்படி,

$$E(\hat{Y}) = E_i \left[E_j \frac{M}{m} \sum_{m} \frac{N_i}{n_i} \sum_{n_i} y_{ij} \right]$$

$j \longrightarrow$ முதல் நிலை கூறினை மாருது வைத்து, இரண்டாம் நிலைக் கூறுகளிடை சராசரி அல்லது எதிர்பார்க்கும் மதிப்பு காணுதல். அதன்பின், எல்லா முதல் நிலைக் கூறுகளையும் கருதி, எதிர் பார்க்கும் மதிப்பினை கணக்கிடுவதை E_i என்பதன் மூலம் குறிக்கிறோம்.

அதாவது (A, B), (B, C), (A, C)

என்ற மூன்று திரள்களின் சேர்வுகளையும் கருதி, அவை ஒவ்வொன்றிலும் உள்ள 9 மாதிரிகளையும் கருதி, சராசரி காண வேண்டும்.

உதாரணமாக, (A, B) என்ற சேர்வுத் திரளில்

$$E_i \frac{M}{m} \sum_{m} \frac{N_i}{n_i} \sum y_{ij}$$

$$= \sum \frac{1}{9} \left[\frac{M}{m} \sum_{m} \frac{N_i}{n_i} \sum^{n_i} y_{ij} \right]$$

{ Σ எல்லா 9 மாதிரிகளையும் கணக்கில் கொள்வது. }

அதாவது, $\frac{1}{9}[28{\cdot}25 + \ldots\ldots + 42{\cdot}75]$

$$= \frac{1}{9} \times (336{\cdot}5)$$

$$= 37{\cdot}4$$

இதேபோல, மற்ற இரு சேர்வுத் திரள்கட்கும் கருதி,

$$E(\hat{Y}) = \frac{1}{3}\left[\frac{1}{9}(28{\cdot}25 + \ldots\ldots + 42{\cdot}75)\right] + \frac{1}{3}\left[\frac{1}{9}(38{\cdot}25 + \ldots\ldots + 54{\cdot}00)\right] + \frac{1}{3}\left[\frac{1}{9}(45{\cdot}00 + \ldots\ldots + 65{\cdot}25)\right]$$

$$= \frac{1}{3 \times 9}[336{\cdot}5 + 459{\cdot}0 + 526{\cdot}0]$$

$$= \frac{1321{\cdot}5}{27} = 49 \text{ ஆகும்.}$$

$Y = 49$ (முழுமைத் தொகுதியினின்று)

$\therefore\ E(\hat{Y}) = Y$ என்று பெறலாம்.

எனவே $\hat{Y}$ ஆனது பிறழ்ச்சியற்றதென அறியலாம்.

(ii) $\hat{Y}$-ன் மாறுபாடு : $[V(\hat{Y})]$

$$S_b^2 = \frac{1}{M-1} \sum_{M} (Y_i - \bar{Y})^2 \qquad \left[\bar{Y} = \frac{Y}{M} = 16{\cdot}3.\right.$$

$$= \frac{1}{2}[(10 - 16{\cdot}3)^2 + (15 - 16{\cdot}3)^2 + (24 - 16{\cdot}3)^2]$$

$$= \frac{100 \cdot 67}{2} = 50 \cdot 335$$

$$S_i^2 = \frac{1}{M-1} \sum_{N_1} (y_{1i} - \bar{Y}_j)^2$$

$$= \frac{1}{(3-1)} [(2 - 3 \cdot 3)^2 + (3 - 3 \cdot 3)^2 + (5 - 3 \cdot 3)^2]$$

$$= \frac{4 \cdot 67}{2} = 2 \cdot 335.$$

இதேபோல்

$$S_2^2 = 8 \; ; \quad S_3^2 = 13.$$

$$\therefore \quad V(\hat{Y}) = \left[3^2 . \left(\frac{3-2}{3} \right) . \frac{50 \cdot 335}{2} \right]$$

$$+ \frac{3}{2} \left[3^2 . \left(\frac{3-2}{3} \right) \left(\frac{1}{2} \right) (2 \cdot 3 + 8 + 13) \right]$$

$$= 127 \cdot 93.$$

எடுத்துக்காட்டு 2 :

முழுமைத் தொகுதி பின்வருமாறு :

திரள்	Y_{ij}
A	1, 6, 14
B	3, 3. 9
C	5, 6, 7

பின்வருவனவற்றைக் கணக்கிடுகிறோம்.

(i) முதலில் $m = 2$ என்றும், $n' = 2$ என்றும் மாதிரிகளை தெரிவு செய்தல் :

மொத்தம் $\binom{M}{m}\binom{N_1}{n_1}\binom{N_2}{n_2}$ $(n_1 = n_2 = n' = 2)$

அதாவது $\binom{3}{2}\binom{3}{2}\binom{3}{2} = 27$ வழிகளில் மாதிரிகள் எடுக்கப்படும்.

(ii) உதாரணமாக,

A : (1, 6) ; B = (3, 3)

என்ற மாதிரிச் சேர்வைக் (Combination) கருதுவோம்.

இதில்,

$$\hat{Y} = \frac{M}{m} \sum_{m} \frac{N_i}{n_i} \sum_{n_i} y_{ij} \qquad [M = 3. \quad [N_1 = N_2 = 3$$

$$= \frac{3}{2}\left[\frac{3}{2}(1+6) + \frac{3}{2}(3+3)\right]$$

$$\therefore \quad \hat{Y} = 29{\cdot}25.$$

(iii) $\hat{V}(\hat{Y})$ மதிப்பு :

முதலில், s_b^2, s_1^2, s_2^2 இவற்றை அறிதல் :

$$s_b^2 = \frac{1}{m-1} \sum_{m} (\hat{Y}_i - \hat{Y})^2$$

$$= [(10{\cdot}5 - 9{\cdot}75)^2 + (9 - 9{\cdot}75)^2] = 1{\cdot}125$$

$$\hat{Y}_1 = N_1 \cdot \bar{y}_1 = 10{\cdot}5, \quad \hat{Y}_2 = N_2 \cdot \bar{y}_2 = 9. \quad \hat{Y} = \frac{1}{m} \Sigma \hat{Y}_i = 9{\cdot}75$$

$$s_1^2 = \frac{1}{(n_1 - 1)} \sum_{n_1} (y_{1j} - \bar{y}_1)^2$$

$$= \frac{1}{(2-1)} [(1 - 3{\cdot}5)^2 + (6 - 3{\cdot}5)^2] = 12{\cdot}5$$

$$s^2_2 = \frac{1}{(n_2 - 1)} \sum_{n_2} (y_{2j} - \bar{y}_2)^2$$

$$= [(3-3)^2 + (3-3)^2] = 0$$

$$\therefore \quad \hat{V}(\hat{Y}) = M^2 . \frac{M-m}{M} . \frac{s_b^2}{m}$$

$$+ \frac{M}{m} \sum_{m} N_i^2 . \frac{N_i - n_i}{N_i} . \frac{s_i^2}{n_i} .$$

$$= \left[3^2 . \left(\frac{1}{3}\right) . \frac{(1 \cdot 125)}{2} \right] + \frac{3}{2} \left[3^2 . \frac{1}{3 \times 2} (12 \cdot 5) \right]$$

$$= \frac{3}{2} . (1 \cdot 125) + \frac{9}{4} (12 \cdot 5)$$

$$\therefore \quad \hat{V}(\hat{Y}) = 29 \cdot 8 .$$

இதேபோல, ஏனைய 26 மாதிரிகளுக்கும் $\hat{V}(\hat{Y})$ மதிப்புகளை கணக்கிட்டு, இறுதியில் $E[\hat{V}(\hat{Y})] = V(\hat{Y})$ என்று நிரூபிக்கலாம்.

மாதிரிகள் :

	A	B		A	C		B	C
1.	1, 6	3, 3	10.	1, 6	5, 6	19.	3, 3	5, 6
2.		3, 9	11.		5, 7	20.		5, 7
3.		3, 9	12.		6, 7	21.		6, 7
4.	6, 14	3, 3	13.	6, 14	5, 6	22.	3, 9	5, 6
5.		3, 9	14.		5, 7	23.		5, 7
6.		3, 9	15.		6, 7	24.		6, 7
7.	1, 14	3, 3	16.	1, 14	5, 6	25.	3, 9	5, 6
8.		3, 9	17.		5, 7	26.		5, 7
9.		3, 9	18.		6, 7	27.		6, 7

s_b^2, s_1^2, s_2^2, $\hat{V}$ $(\hat{Y})$ ஆகியவற்றின் அட்டவணை :

மாதிரி எண்	s_b^2	s_1^2	s_2^2	$\hat{V}$
1.	1·125	0·125	0	2·569
2.	28·125	0·125	18·0	82·969
3.	28·125	0·125	18·0	82 969
4.	220·50	32·0	0	402·75
5.	72·0	32 0	18·0	220·5
6.	72·0	32·0	18·0	220·5
7.	91·125	84·5	0	326·813
8.	10·125	84·5	18·0	245·813
9.	10·125	84·5	18·0	245·813
10.	18·0	12·5	0·5	56·25
11.	28·125	12·5	2·0	59·813
12.	40·50	12·5	0·5	90·0
13.	91·125	32·0	0·5	209·813
14.	72·0	32·0	2·0	184·5
15.	55·125	32·0	0·5	155·813
16.	18·0	84·5	0·5	218·25
17.	10·125	84·5	2·0	209·813
18.	4·5	84·5	0·5	198·0
19.	21·125	0	0 5	32·813
20.	40·5	0	2·0	65·25
21.	55·125	0	0·5	166 50
22.	1·125	18·0	0·5	43·313
23.	0	18·0	2·0	45·0
24.	1·125	18·0	0·5	43·313
25.	1·125	18·0	0·5	43·313
26.	0	18·0	2·0	45·0
27.	1·125	18·0	0·5	43·313

$$\therefore \quad E[\hat{V}(\hat{Y})] = \frac{1}{27} \Sigma \hat{V} (\hat{Y})$$

$$= \frac{3740 \cdot 763}{27} = 138 \cdot 55$$

$$\therefore \quad E[\hat{V}(\hat{Y})] \simeq 139.$$

முழுமைத் தொகுதியில் பின்வருமாறு $V(\hat{Y})$ பெறப்படும்.

Y_{ij}	Y_i	$\bar{Y}_i$
1, 6, 14	21	7
3, 3, 9	15	5
5, 6, 7	18	6

$$\therefore \quad \bar{Y} = \frac{Y}{M} = \frac{54}{3} = 18.$$

$\hat{Y}$-ன் மாறுபாட்டு அளவு :

$$V(\hat{Y}) = \left[M^2 \cdot \left(\frac{M-m}{M}\right) \cdot \frac{S_b^2}{m}\right] + \frac{M}{m} \sum_{1}^{M} N_i^2 \cdot \frac{N_i - n_i}{N_i} \cdot \frac{S_i^2}{n_i}.$$

இதற்கு,

$$S_b^2 = \frac{1}{M-1} \sum_{M} (Y_i - \bar{Y})^2$$

$$= \frac{1}{2} [(21-18)^2 + (15-18)^2 + (18-18)^2]$$

$$\therefore \quad S_b^2 = \frac{18}{2} = 9$$ எனவும்,

$$S_1^2 = \frac{1}{N_1 - 1} \sum_{N_1} (Y_{1j} - \bar{Y}_1)^2$$

$$= \frac{1}{2} [(1-7)^2 + (6-7)^2 + (14-7)^2]$$

$$= \frac{1}{2} [36 + 1 + 49] = \frac{86}{2}$$

$S_1^2 = 43$ எனவும், இதேபோல்

$S_2^2 = 12$; $S_3^2 = 1$ எனவும் பெறப்படுகின்றன.

$$\therefore V(\hat{Y}) = \left[3^2 \left(\frac{3-2}{3} \right) \times \frac{9}{2} \right] + \frac{3}{2} \left[\frac{3}{2} (43 + 12 + 1) \right]$$

$$\therefore V(\hat{Y}) = 13{\cdot}5 + 126 = 139{\cdot}5$$

இதிலிருந்து $E\left[\hat{V}(\hat{Y}) \right] = V(Y)$ என்று அறியலாம்.

எடுத்துக்காட்டு 3 :—

திரள்	Y_{ij}	Y_i	$\bar{Y}_i$
A	1, 3, 5	9	3
B	3, 5, 7	15	5
C	5, 7, 9	21	7

ஒரு மாதிரியினைப் பின்வருமாறு தெரிவு செய்வோம் :

A : (3, 5) ; C : (7, 9)

y_{1j}	y_{2j}	y_{1j}^2	y_{2j}^2
3	7	9	49
5	9	25	81
8	16	34	130

$$\text{(i)} \quad \hat{Y} = \frac{1}{N} \frac{M}{m} \sum_{m} \frac{N_i}{n_j} \sum_{n_i} y_{ij}$$

$$= \frac{1}{9} \frac{3}{2} \left[\frac{3}{2} (3+5) + \frac{3}{2} (7+9) \right]$$

$$\hat{\bar{Y}} = \frac{1}{4} (24) = 6$$

$$\text{(ii)} \quad \hat{V}(\hat{\bar{Y}}) = \frac{M-m}{M} \frac{s_{1b}^2}{m} + \frac{N'-n'}{N'} \frac{1}{M} \sum_{m} \frac{s_i^2}{mn'}$$

$M = 3$; $N' = 3$; $n' = 2$; $m = 2$.

$s_{ib}^2 = \frac{1}{m-1} \left(\sum \bar{y}_i^2 - m\bar{y}^2 \right)$ ஆகும்.

இதில்,

$$\sum_{m} \bar{y}_i^2 = \left(\frac{1}{n'}\right)^2 \sum_{m} \left(\sum_{n} y_{il}\right)^2$$

$$= \left(\frac{1}{n'}\right)^2 \left[\left(\sum_{n'} y_{1j}\right)^2 + \left(\sum_{n'} y_{2j}\right)^2 \right]$$

$$\therefore \sum_{m} \bar{y}_i^2 = \left(\frac{1}{2}\right)^2 \left[8^2 + 16^2 \right]$$

$$\sum_{m} \bar{y}_j^2 = \frac{320}{4} = 80$$

$$\bar{y}^2 = \left(\frac{1}{mn'}\right)^2 \left(\sum_{m} \sum_{n'} y_{ij}\right)^2$$

$$= \frac{1}{(2 \times 2)^2} \left(\Sigma y_{ij} + \Sigma y_{2j} \right)^2$$

$$\therefore \bar{y}^2 = \frac{1}{16} (8 + 16)^2 = \frac{24 \times 24}{16} = 36$$

எனவே,

$$S_{1b}^2 = \frac{1}{m-1} \left(\Sigma \bar{y}_i^2 - m\bar{y}^2 \right)$$

$$= \frac{1}{(2-1)} [80 - 72] = 8.$$

அடுத்து,

$$s_{2i}^2 = \frac{1}{m(n-1)} \sum_{m} \sum_{n'} y_{ij} - \bar{y}_i)^2$$

$$= \frac{1}{m(n'-1)} \left[\sum_{m} \sum_{n'} y_{ij}^2 \frac{1}{n'} \sum_{m} \left(\sum_{n'} \bar{y}_{ij}\right)^2 \right]$$

$$= \frac{1}{2(2-1)} \left\{ (34 + 130) - \frac{1}{2}\left(8^2+16^2\right) \right\}$$

$$\therefore \quad S_{2i}^2 = \frac{1}{2} \; [164 - 160] = 2$$

$$\hat{V}_{(\bar{Y})} = \frac{M-m}{M} \cdot \frac{S_{1b}^2}{m} + \frac{N^1-n^1}{N'} \cdot \frac{m}{M} \cdot \frac{S_{2i}^2}{mn'}$$

$$= \frac{3-2}{3} \cdot \frac{8}{2} + \frac{3-2}{3} \cdot \frac{2}{3} \cdot \frac{2}{2\times 2}$$

$$\therefore \hat{V}\left(\bar{Y}\right) = \frac{4}{3} + \frac{1}{9} = \frac{13}{9}$$

எடுத்துக்காட்டு : 4.

முழுமைத் தொகுதி :

திரள்	Y_{ij}
A	1, 3, 5.
B	3, 5, 7.
C	5, 7, 2.

(1) $\sum_{1}^{M} \sum_{f<j'}^{N'} (Y_{ij}-\bar{Y})(Y_{ij'}-\bar{Y})$ **மதிப்பினை கணக்கிடல் :**

$M = 3 ; \quad N' = 3.$

$\bar{Y} = \frac{Y}{MN'}$ —(வரையறை)

$\therefore \quad \bar{Y} = \frac{45}{9}$ **= 5 என்று அறியப்படுகிறது.**

எனவே,

$$\sum_{M} \sum_{j<j} (Y_{ij} - \bar{Y})(Y_{ij'} - \bar{Y})$$

$$= [\{(-4)\times(-2) + (-4)\times 0 + (-2)(0)\}$$
$$+ \{(-2\times 0 + (-2\times 2) + (0\times 2)\}$$
$$+ \{(0\times 2) + (0\times 4) + (2\times 4)\}$$

$$= 8 - 4 + 8$$

$$= 12$$

என்று பெறுகிறோம்.

(ii) மாறுபாடு :

$$S^2 = \frac{1}{MN' - 1} \sum_{1}^{M} \sum_{1}^{N'} (Y_{ij} - \bar{Y})^2$$

$$= \frac{1}{8} (16 + 4 + 0 + 4 + 0 + 4 + 0 + 4 + 16)$$

$$= \frac{48}{8} = 6$$

$\therefore \quad S^2 = 6$ என்று வரும்.

(iii) உள்ளீட்டு ஒட்டுறவுக் கெழு :

$$\rho = \frac{E\,(Y_{ij} - \bar{Y})\,(Y_{ij}' - \bar{Y})}{E\,(Y_{ij} - \bar{Y})^2}$$

$$= \frac{2 \sum_{M} \sum_{N'} (Y_{ij} - \bar{Y})\,(Y_{ij}' - \bar{Y})}{(N' - 1)\,(MN' - 1) \cdot S^2}$$

$\therefore$ (i), (ii) இவற்றைப் பயன்படுத்தி,

$$\rho = \frac{2 \times 12}{2 \times 8 \times 6} = \frac{3}{12} = 0{\cdot}25$$ என்று பெறுகிறோம்.

(iv) மேற்கண்ட முடிவைப் பயன்படுத்தி,

$$V(\bar{Y}_{cl}) = \frac{(MN' - 1)}{MN'} \cdot \frac{S^2}{mn'} \cdot \left[\frac{N' - n'}{N'} + \frac{M - m}{M - 1} \cdot \frac{n'}{N'} \right.$$

$$\left. + \left\{ \frac{M - m}{M - 1} \cdot \frac{n'}{N'} \cdot (N' - 1) - \frac{N' - n'}{N'} \right\} \cdot \rho \right]$$

{ $m = 2$; $n' = 2$ என்ற அளவுகளில் ஒவ்வொரு நிலையிலும் முறையே மாதிரிகள் எடுக்கப்படுவதாகக் கொள்க}.

$$\therefore \quad V(\hat{\hat{Y}}_{cl}) = \frac{8}{9} \cdot \frac{6}{4} \cdot \left[\frac{1}{3} + \left(\frac{1}{2} \cdot \frac{2}{3}\right) + \left(\frac{1}{3} \cdot \frac{\ }{12}\right)\right]$$

$$= \frac{8}{9} \cdot \frac{6}{4} \cdot \left[\frac{9}{12}\right]$$

$$\therefore \quad V(\hat{\hat{Y}}_{cl}) = 1$$

என்றபடி, $\hat{\hat{Y}}$-ன் மாறுபாட்டு அளவு பெறப்படுகிறது.

(v) மொத்த மதிப்பீட்டின் மாறுபாடு :

அதாவது $V(\hat{Y})$ மதிப்பிடல் ;

இதற்கு முறையே S_b^2, S_1^2, S_2^2, S_3^2 ஆகியவற்றைக் கணக்கிட்டுப் பெறவேண்டும்.

வரையறைப்படி,

$$S_b^2 = \frac{1}{M-1} \sum_{M} (Y_i - \bar{Y})^2 \text{ ஆகும்.}$$

$$\bar{Y} = \frac{1}{M} \Sigma\, \Sigma\, Y_{ij} = \frac{45}{3} = 15 \text{ என்பதாகும்.}$$

$$\therefore \quad S_b^2 = \frac{1}{2} \left[(9-15)^2 + (15-15)^2 + (21-15)^2 \right]$$

$$\therefore \quad S_b^2 = \frac{1}{2} \cdot (72) = 36$$

மேலும்,

$$S_i^2 = \frac{1}{(N_i - 1)} \sum_{j=1}^{N_i} (Y_{ij} - \bar{Y}_i)^2$$

என்பதில் $i = 1, 2, 3$ என்று பிரதியிட, S_1^2, S_2^2, S_3^2 என்று பெறப்படும்.

$$\therefore \quad S_1^2 = \frac{1}{2}[(1-3)^2 + (3-3)^2 + (5+3)^2]$$

$$= \frac{8}{2} = 4 \text{ (முதல் திரள்)}$$

$S_2^2 = 4$; $S_3^2 = 4$ (என்று காணலாம்).

எனவே,

$$V(\hat{Y}) = M^2 . \frac{M-m}{M} . \frac{S_b^2}{m} + \frac{M}{m} \Sigma N_i^2 . \frac{N_i - n_i}{N_i} . \frac{S_i^2}{n_i}$$

$$= \left[9 \cdot \left(\frac{1}{3}\right)(18)\right] + \frac{3}{2}\left[9 . \frac{1}{3} . \frac{1}{2}(4 + 4 + 4)\right]$$

$[m = 2, n_i = n' = 2] \ (N_i = N' = 3)$

$$\therefore \quad V(\hat{Y}) = 54 + 27 = 81.$$

(vi) S_{1b}^2-ன் மதிப்பை அறிதல் :

வரையறை :

$$S_{1b}^2 = \frac{1}{M-1} \sum_M (\bar{Y}_i - \bar{Y})^2$$

$$= \frac{1}{N' . (M-1)} \sum_M (Y_i - \bar{Y})^2$$

$$= \frac{(MN' - 1)}{(N')^2 . (M-1)} . S^2 [1 + (N'-1)\rho]$$

$$= \frac{8}{9 \times 2} . 6 \left[1 + \left(2 \times \frac{3}{12}\right)\right]$$

$$= \frac{4 \times 2}{3} . \left[\frac{3}{2}\right]$$

$$\therefore \quad S_{1b}^2 = 4$$

என்று பெறப்படும்.

(vii) S_{2i}^2-ன் மதிப்பு :

வரையறை :

$$S_{2i}^2 = \frac{MN'-1}{MN'} . \ S^2 \ (1-\rho)$$

$$= \frac{8}{9} . 6 \left(1 - \frac{3}{12}\right)$$

$$\therefore \quad S_{2i}^2 = 4$$

என்று எளிதாகக் கணக்கிடுகிறோம்.

பயிற்சிகள்

1. பின்வரும் விவரங்கள் கொடுக்கப்படுகின்றன. இவற்றில், மூன்று பிரிவுகளில் உள்ள மாணவர்கள் வைத்திருக்கும் புத்தகங்கள் எண்ணிக்கைகள் த்ரப்படுகின்றன.

திரள்	Y_{ij} (புத்தகங்கள்)
A	10, 30, 50
B	15, 25, 35
C	25, 35, 45

இவற்றிலிருந்து இரு பிரிவுகளை ராண்டம் அடிப்படையில் (ஈடு செய்யப்படாத முறைப்படி) தெரிவு செய்து, அவை ஒன்றிலும் இரு கூறுகளை (subunits) தெரிவு செய்து, பின் வருவனவற்றை மதிப்பிடுக.

(i) Y-ன் மதிப்பீட்டைக் கணக்கிடுக.

(ii) எல்லா மாதிரிகளையும் கருதி, $\hat{Y}$ ஆனது Y-ன் பிறழ்ச்சியற்ற மதிப்பீடாக அமைவதெனக் காட்டுக.

2. தெரிந்தெடுக்கப்பட்ட நான்கு கிராமங்களில், மாதிரியாக எடுக்கப்பட்ட பண்ணைகளில் உள்ள கால்நடை விவரங்களாவன : (N_i : ஒவ்வொரு கிராமத்திலும் மொத்த பண்ணைகள்).

கிராமம்	N_i	y_{ij}
1	20	12, 15, 18, 19, 11
2	35	13, 10, 17, 16, 14
3	30	16, 17, 15, 12
4	25	14, 19, 13, 11, 12, 15

y_{ij} என்பது i-ஆவது கிராமத்தில் உள்ள j-ஆவது பண்ணையில் உள்ள கால்நடைகள் எண்ணிக்கை. மொத்த கால்நடைகளின் எண்ணிக்கையான Y-க்கு மதிப்பீட்டினைக் காண்க. ($M = 20$ என்று கொள்க.)

$\bar{Y}$-ன் மதிப்பீடு என்ன ?

3. ஒரு நிறுவனமானது, ஒரு வாரத்தில் தன்னிடம் வருகை தரும் பார்வையாளர்களைப் பற்றிய மதிப்பீடு செய்ய எண்ணிப் பின்வரும் விவரங்களைப் பெறுகிறது. அதனுள் மொத்தமாக 50 பிரிவுகள் உள்ளன. அவற்றிலிருந்து ஐந்தை தேர்ந்தெடுத்து, அவற்றில் உள்ள பணியாளரிடம் விவரங்கள் சேகரிக்கப் படுகின்றன. ஒவ்வொரு பிரிவிலும் N_i பணியாளர்கள் பணியாற்றுகின்றனர். தெரிந்தெடுக்கப்பட்ட 5 பிரிவுகளில் உள்ள பணியாளர் தம்மைக் காண வந்தோர் பற்றி குறித்து வைக்கும் விவரங்கள் தரப்படுகின்றன.

பிரிவு	N_i	y_{ij} (பார்வையாளர் எண்ணிக்கை)
1	10	5, 3, 0, 8
2	15	11, 6, 5, 9, 4
3	10	4, 3, 5
4	20	6, 8, 2, 4
5	15	2, 0, 9, 9

(i) மொத்த பார்வையாளர்கள் எண்ணிக்கையை மதிப்பீடு செய்க.

(ii) s_i^2, s_b^2 ஆகியவற்றை கணக்கிடுக.

(iii) இவற்றின் மூலம், $\hat{V}(\hat{Y})$-ன் மதிப்பை அறிக.

(iv) $s(\hat{Y})$-ன் மதிப்பினையும் தருக.

4.

திரள்	Y_{ij}
A	10, 30, 50
B	30, 50, 70
C	50, 70, 90

$m = 2$ என்று கொண்டு, ஒவ்வொரு தெரிந்தெடுக்கப்பட்ட திரளினின்றும் இரு துணைக் கூறுகள் எடுப்பதாகக் கொண்டு

(i) $\hat{Y}$, $\hat{\hat{Y}}$ ஆகியவற்றினைக் காண்க.

(ii) s_b^2, s_i^2 $(i = 1, 2)$ ஆகியவற்றை மதிப்பிடுக.

(iii) எல்லா மாதிரிகளையும் கருதி,

$$E[V(\hat{Y})] = V(\hat{Y})$$ என்று பெறுவதை நிறுவுக.

(குறிப்பு : மொத்தம் 27 மாதிரிகள் அமைகின்றன.)

5. பின்வரும் விவரங்கள் கொடுக்கப்படுகின்றன :

பிரிவு	Y_{ij}
A	1, 6, 14
B	3, 3, 9
C	5, 6, 7

(i) $V(\hat{Y})$-ன் மதிப்பினைக் கணக்கிடுக.

(ii) $m = 2$ என்றும், $n' = 2$ என்றும் கருதி எடுக்கப்படும் மாதிரிகளை பட்டியலில் அமைத்துக் காட்டுக.

(iii) இவை ஒவ்வொன்றிற்கும் s_1^2, s_2^2, s_b^2 இவற்றின் மதிப்புகளைப் பெறுக.

(iv) $E(s_1^2) = S_1^2$ என்று காட்டுக.

(v) மேலும் $E(s_b^2) = S_b^2 + \frac{1}{M} \sum_{M} N_i^2 \cdot \frac{N_i - n_i}{N_i} \cdot \frac{S_i^2}{n_i}$

என்பதனை நிறுவுக.

6. பின்வரும் முழுமைத் தொகுதியானது கொடுக்கப்படுகிறது :

திரள்	Y_{ij}
A	1, 3, 5
B	3, 6, 6
C	7, 9, 14

$m = 3$ என்று முதல்நிலைக் கூறுகளையும், ஒவ்வொரு திரளினின்றும் $n' = 2$ என்று துணைக் கூறுகளையும் தெரிவு செய்க.

(i) எடுக்கப்பட்ட மாதிரியினைப் பயன்படுத்தி $\hat{\hat{Y}}$-ன் மதிப்பைக் காண்க.

(ii) S_{1b}^2 S_{2i}^2 இவற்றினைக் கணக்கிடுக.

(iii) தெரிவு செய்யப்பட்ட மாதிரியினைக் கொண்டு, s_{1b}^2 மற்றும் s_{2i}^2 இவற்றினை மதிப்பிடுக.

(iv) $V(\hat{\hat{Y}})$-ன் மதிப்பு என்ன ?

7. கருதப்பட்ட பிரிவினர் சராசரியாகச் செலவு செய்யும் பண விவரம் (ஒரு மாதத்தினைக் கருதி) பின்வருமாறு : $M = 400$ என்றும், $N' = 40$ என்றும் கொண்டு, அதனின்று $m = 10$ எனவும், $n' = 10$ எனவும் தெரிவு செய்யப்பட்ட மாதிரிகளின் மதிப்புகள் கொடுக்கப்படுகின்றன : (ரூபாய்களில்)

பிரிவு	$\sum_j y_{ij}$	பிரிவு	$\sum_j y_{ij}$
1	250	6	180
2	300	7	240
3	200	8	270
4	190	9	200
5	230	10	210

(i) $\hat{\hat{Y}}$-ன் மதிப்பினைக் காண்க.

(ii) $\hat{V}(\hat{\hat{Y}}) = \frac{s_{1b}^2}{m} = \frac{1}{m(m-1)} \sum_m (\bar{y}_i - \bar{\bar{y}})^2$

என்பதனைப் பயன்படுத்தி $\hat{V}(\hat{\hat{Y}})$-ன் மதிப்பினை கணக்கிடுக.

8. பின்வரும் முழுமைத் தொகுதியினைக் கருதி,

(i) ρ-மதிப்பினைக் கணக்கிடுக.

(ii) $V(\hat{Y}_{cl})$ மதிப்பினை, ρ வினை பயன்படுத்திப் பெறுக.

(iii) S_{1b}^2, S_{2i}^2 ஆகியவற்றினையும், ρ-மூலம் பெறுக.

திரள்	Y_{ij}
A	1, 1, 1
B	2, 2, 2
C	3, 3, 3

(vii) திரள் மாதிரி முறை, எளிய ராண்டம் மாதிரிமுறை பற்றிய ஒப்புமை : (Comparison)

சிறந்த மாதிரியினைத் திட்டமிடுதலில், குறைந்த அளவு மாறுபாடுடைய மதிப்பீடுகளைப் பெறுவதே அடிப்படை நோக்கம். வெவ்வேறு திரள்களின் உள்ளீட்டு மாறுபாடு குறைவானதாக இருப்பின், சிறிய அளவில் இரண்டாம் கட்டத் தெரிவு அமைதல் நன்று. ஒரு பொருளின் தரமானது ஒரே மாதிரித் தன்மை உடையதாகக் கொண்டால், அதை உற்பத்தி செய்யும் முறைகளில் (process) பல்வகையான மாறுபாடுகள் இருக்கலாம். இந்நிலையில், துணைக்கூறுகளை சிறிய அளவிலும், முதற்கட்டத் தேர்வினை அதிக அளவிலும் செய்தல் பயன்தரும். அதாவது, திரள்களின் உள்ளீட்டு வேறுபாட்டினைப் பொறுத்து, எடுக்கப்படும் துணைக்கூறுகளின் அளவு அமைந்திருக்கும்.

இதற்கான வழிமுறை, திரள்களின் உள்ளீட்டு ஒட்டுறவுக் கெழுவின் அடிப்படையில் வகுக்கப்படுகிறது. திரள்முறையின் திறன் அளவு, சமவாய்ப்பு முறையுடன் ஒப்பிடுகையில், இதைப் பொறுத்தே அமையும்.

அதாவது $V(\hat{Y}_c)$ மற்றும் V_{ran} ஆகியவை ஒப்பிடப்பட வேண்டும். முதலில் எளிய சமவாய்ப்பு முறையில் பெறும் $V(\hat{Y}_{ran})$ மதிப்பை காண்போம்.

$$N_i = N' = \frac{N}{M}, \quad n_i = n' = \frac{n}{m} \text{ எனக் கொள்க.}$$

$$\therefore \hat{\bar{Y}} = \frac{1}{mn'} \cdot \sum_{m} \sum_{n'} y_{ij} \text{ ஆகும்.}$$

இது $n = mn'$ மாதிரிக் கூறுகளின் மாதிரிச் சராசரி.

மேலும், $E(\hat{\bar{\bar{Y}}}) = \bar{\bar{Y}}$ ஆகும்.

∴ சமவாய்ப்பு முறைப்படி,

$$V(\hat{Y}_{ran}) = \frac{MN' - mn'}{MN'} \cdot \frac{S^2}{mn'} \quad \ldots\ldots \text{(A)}$$

இதில்,

$$S^2 = \frac{1}{MN' - 1} \sum_{M} \sum_{N'} (Y_{ij} - \bar{Y})^2$$

$$\left\{ \text{இங்கு } \bar{Y} = \frac{Y}{MN'} \cdot \right\}$$

துணைத் தொகுதி அளவு MN' ஆகும்; மாதிரி அளவு mn'. திரள் மாதிரி முறைப்படி,

$$V(\hat{\bar{\bar{Y}}}_c) = \frac{M - m}{M} \cdot \frac{S_{1b}^2}{m} + \frac{N' - n'}{N'} \cdot \frac{S_{2i}^2}{m} \quad \ldots\ldots \text{(B)}$$

(A), (B) இவற்றை ஒப்பிட, S_{1b}^2 ஐயும், S_{2i}^2 ஐயும் S^2 மூலமாக எழுதவேண்டும். பிரிவு உள்ளீட்டு ஒட்டுறவுக் கெழு (intra class correlation coefficient) வான ρ வினைப் பயன் படுத்தி, இடைப்பட்ட (between) மாறுபாட்டை,

$$S_{1b}^2 = \frac{MN' - 1}{(N')^2 (M - 1)} \cdot S^2 [1 + (N' - 1)\rho]$$

என்று எழுதலாம்.

இதில்,

$$\rho = \frac{E(y_{ij} - \bar{Y})(y_{ij}' - \bar{Y})}{E(y_{ij} - \bar{Y})^2}$$

$$= \frac{2 \sum_{i=1}^{M} \sum_{j<k}^{N'} (y_{ij} - \bar{Y})(y_{ij}' - \bar{Y})}{(N' - 1)(MN' - 1)S^2}$$

ஆகிறது.

தேற்றம் :

திரள் மாதிரிமுறையில் வரையறுக்கப்படும்

$$S_{1b}^2 = \frac{1}{M-1} \sum_M (\bar{Y}_i - \bar{Y})^2$$ என்ற, திரள்களின்

இடைப்பட்ட மாறுபாட்டு அளவினை (between clusters),

$$\rho = \frac{E(y_{ij} - \bar{Y})(y_{ij'} - \bar{Y})}{E(y_{ij} - \bar{Y})^2}$$

என்ற திரள்-உள்ளிட்ட ஒட்டுறவுக் கெழுவினை (intracluster correlation coeffcient) பயன்படுத்தி,

$$S_{1b}^2 = \frac{MN' - 1}{(N')^2 (M-1)} \cdot S^2 [1 + (N'-1)\rho]$$

என்று எழுதலாம். S^2 என்பது, முழுமைத்தொகுதியின் மொத்த மாறுபாடாக கருதப்படும். $\left(N' = \frac{N}{M}\right)$. (total variance)

நிரூபணம் :

வரையறைப்படி,

$$S_{1b}^2 = \frac{1}{M-1} \sum_M (\bar{Y}_i - \bar{Y})^2$$

அதாவது,

$$S_{1b}^2 = \frac{1}{(N')^2 (M-1)} \sum_M (Y_i - \bar{Y})^2$$ என்றாகும்.

$(Y_i - \bar{Y})$ என்ற வேறு பாட்டை,

$$(Y_i - \bar{Y}) = (y_{i1} - \bar{Y}) + (y_{i2} - \bar{Y}) + \dots + (y_{iN'} - \bar{Y})$$

என்று விரித்து எழுத, இதிலிருந்து

$$\sum_M (Y_i - \bar{Y})^2 = \sum_M \sum_{N'} (y_{ij} - \bar{Y})^2 + \sum_M \sum_{j<j'}^{N'} (y_{ij} - \bar{Y})(y_{ij'} - \bar{Y})$$

என்று பெறப்படும்.

{ இதேபோன்ற ஒரு நிலை, ஒழுங்குமுறையில் காணப் பட்டதை நினைவு கூர்க. }

மேற்கண்ட சமன்பாட்டில், வலப்புற முதல் உறுப்பு, S^2-உடன் தொடர்புற்றிருப்பதையும், இரண்டாம் உறுப்பு, y_{ij}-யின் இணைமதிப்புகளினிடையே (pairs) உள்ள ஒட்டுறவை குறிப்பதையும் கவனிக்க.

திரள் உள்ளிட்ட ஒட்டுறவுக் கெழுவான ρ வின் வரையறை யில், தொகுதியானது (Numerator)

$$E(y_{ij} - \bar{Y})(y_{ij}' - \bar{Y}) = \sum_{i=1}^{M} \sum_{j<j'}^{N'} \frac{1}{[MN'(N'-1)/2]} \cdot$$

$(y_{ij} - \bar{Y})(y_{ij}' - \bar{Y})$ ஆகிறது. [$\because$ (y_{ij}, y_{ij}') எனும் இணைகளை, மொத்தம் $\frac{N'(N'-1)}{2}$ வழிகளில் பெறலாம். இதே போல் M திரள்கட்கு $M[N'(N'-1)/2]$ வழிகள் அமையும்].

ρ வின் பகுதியானது (Denominator),

$$E(y_{ij} - \bar{Y})^2 = \sum_{M} \sum_{N'} \frac{1}{MN'} (y_{ij} - \bar{Y})^2$$

$$= \frac{1}{MN'} (MN' - 1) S^2 \text{ ஆகும்.}$$

எனவே,

$$\rho = \frac{2 \sum_{M} \sum_{j<j'} (y_{ij} - \bar{Y})(y_{ij}' - \bar{Y})}{(N'-1)(MN'-1). S^2}$$

ஆக, இதிலிருந்து,

$$2 \Sigma \Sigma (y_{ij}' - \bar{Y})(y_{ij}' - \bar{Y}) = \rho (N'-1)(MN'-1). S^2$$

எனப்பெற்று, $\sum_{1}^{M} (Y_i - \bar{Y})^2$-ன் மதிப்பில் பிரதியிடலாம்.

எனவே,

$$\sum_{M} (Y_i - \bar{Y})^2 = \sum_{M} \sum_{N'} (y_{ij} - \bar{Y})^2$$

$$+ 2 \sum_{M} \sum_{j<j'} (Y_{ij} - \bar{Y}) (y_{ij}' - \bar{Y})$$

$$= (MN' - 1). \quad S^2$$
$$+ \rho (N' - 1) (MN' - 1). \quad S^2$$

$$= (MN' - 1). \ S^2. \quad [1 + (N' - 1) . \rho]$$

என்றாகும்.

இதன் மூலம்,

$$S_{1b}{}^2 = \frac{1}{(N')^2 (M - 1)} . \sum_{M} (Y_i - \bar{Y})^2$$

$$= \frac{MN' - 1}{(N')^2 (M - 1)} . \ S^2 \ [1 + (N' - 1). \ \rho]$$

என்று நிரூபிக்கப்படும்

தேற்றம் :—

திரள் மாதிரி முறையில், பல்வேறு திரள்களின் (clusters) உள்ளீட்டு மாறுபாடான (within variation)

$$S_{2i}{}^2 = \frac{1}{M} \sum_{M} S_i{}^2$$

$$= \frac{1}{M (N'-1)} \sum_{M} \sum_{N'} (y_{ij} - \bar{y}_i)^2, \left(N' = \frac{N}{M} \right)$$

என்பதனை,

$$S_{2i}{}^2 = \frac{MN'-1}{MN'} \cdot S^2 \cdot (1-\rho)$$

என்று, S^2, ρ ஆகியவற்றின் மூலமாக எழுதலாம்.

நிரூபணம் :

பின்வரும் சமன்பாட்டைக் கருதுக.

$$\sum_{M}\sum_{N'}(y_{ij}-\overline{Y})^2 = N'\sum_{M}(\overline{Y}_i-\overline{Y})^2 + \sum_{M}\sum_{N'}(y_{ij}-\overline{Y}_i)^2$$

இது ஒரு முற்றொருமையாகும் (identity),

இதிலிருந்து, $(MN'-1)\,S^2 = (M-1)\,N'\cdot S_{1b}^2 + M(N'-1)\,S_{2i}^2$ என்பது புலனாகும்.

S_{1b}^2—க்கு முந்திய தேற்றப்படி பிரதியிட

$$M(N'-1)\,S_{2i}^2 = (MN'-1)\,S^2 - (M-1)\,N'\frac{MN'-1}{N'(M-1)}\cdot\frac{S}{N'}\cdot\times[1+(N'-1)\rho]$$

$$= (MN'-1)S^2 - \frac{(MN'-1)}{(N')^2}S^2[1+(N'-1)\cdot\rho]$$

$$= \frac{S^2}{N'}(MN'-1)[N'-1)-(N'-1)\rho]$$

$$\therefore\ S_{2i}^2 = \frac{MN'-1}{MN'}\cdot S^2\cdot(1-\rho)$$

என்று பெறப்படும்.

மேற்கண்ட இரு தேற்றங்களின் முடிவுகள் மூலமாக S_{1b}^2 S_{2i}^2 ஆகியவற்றை, S^2, ρ மூலமாகப் பெற்று அவற்றைத் தக்கபடி பிரதியிட, ஏற்கெனவே வரையப்பட்ட மாறுபாடு (B) ஆனது.

$$V(\hat{Y}_c) = \frac{MN'-1}{MN'}\cdot\frac{S^2}{mn'}\left[\frac{N'-n'}{N'}+\frac{M-m}{M-1}\cdot\frac{n'}{N'} + \left\{\frac{M-m}{M-1}\cdot\frac{n'}{N'}\cdot(N'-1)-\frac{N'-n'}{N'}\right\}\cdot\rho\right]$$

இதில்,

$\frac{MN'-1}{MN'} \cdot \frac{S^2}{mn'}$ என்ற காரணியானது. ஏறக்குறைய

$$V\ (\hat{\bar{Y}}_{ran}) = \frac{MN'-mn'}{MN'} \cdot \frac{S^2}{mn'}$$

என்பதற்குச் சமமாக அமைகிறது.

{ MN' மிகப் பெரிய அளவினதாகவும், mn', மிகவும் சிறியதாகவும் இருக்கும் நிலையில், இவை சமமாக இருக்க வாய்ப்பு உண்டு.}

எனவே சில நிலைகளில் $V\ (\hat{\bar{y}}_c)$ யானது $V\ (\hat{\bar{y}}_{ran})$ மதிப்போடு சமமாக இருப்பதற்கு வாய்ப்பு உண்டு.

n'—உடன் ஒப்பிடுகையில், மிகப் பெரியதாக N'— மதிப்பையும், m—உடன் ஒப்பிடுகையில், மிகப் பெரியதாக M மதிப்பையும் கருதுவோம். அந்நிலையில்

(i) $\frac{N'-n'}{N'} \simeq 1$ [$\simeq$ என்பது 'ஏறக்குறைய' எனக் குறிக்கும்.]

(ii) $\frac{M-m}{M-1} \cdot \frac{n'}{N'} \simeq 0$

(iii) $\frac{M-m}{M-1}\ \frac{n'}{N'} \cdot (N'-1) \simeq n'$

என்று பெறப்படுவதால்,

$$V\ (\hat{\bar{y}}_c) = \frac{MN'-1}{MN'} \cdot \frac{S^2}{mn'} (1 + 0 + n' - 1\ \rho)$$

$$\simeq \frac{S^2}{mn'} (1 + \overline{n'-1} \cdot \rho)$$

ஆனால்,

$$V_{ran} = \frac{S^2}{mn'} \text{ என்பதால்,}$$

$$V_c = V_{ran} (1 + \overline{n'-1} \cdot \rho) \text{ ஆகும்.}$$

(V_c : திரள் மாதிரியில் மாறுபாட்டு அளவு ;

V_{ran} ராண்டம் மாதிரியில் மாறுபாட்டு அளவு)

முடிவு :

எனவே, $mn' = n$ ஆனால், அதாவது மாதிரியின் அளவு மாறிலியாக அமைந்தால் (constant) மேலும், $\rho > 0$ எனவும் அமையப் பெறின், n' அதிகரிக்க அதிகரிக்க, V_c யானது அதிக அளவில் V_{ran} உடன் ஒப்பிடப்படுகிறது.

$$V(\hat{Y}_c) = \frac{S^2}{mn'} (1 + n' - 1 \cdot \rho)$$

என்ற வரையறையில்,

(i) S^2 என்ற பகுதி எப்போதும் மாருத ஒன்ருகும்.

(ii) m, n', ρ ஆகியவை மாறுவதற்கேற்ப $V(\hat{Y}_c)$ யானது பலவிதத்தில் மாறுபடுகிறது.

முதல்நிலை [Case (i)] : m, n' என்பவை மாறுகையில், $mn' = n =$ மாறிலி எனக் கொள்ளல் வேண்டும். உதாரணமாக, $mn' = 400$ எனில்,

$n' = 2$ என்ற நிலையில், $m = 200$ என்றும்,

$n' = 4$ என்ற நிலையில், $m = 100$ என்றும்,

$n' = 10$ என்ற நிலையில், $m = 40$ என்றும்,

இதேபோன்று விகிதத்தில் மதிப்புகள் பெறப்படுகின்றன. இவை ஏதேனும் ஒரு நிலையில், $\frac{S^2}{mn'}$ மாருத மதிப்புடன் அமையும்.

இரண்டாம் நிலை [case (ii)] : n' மாறுபடுகையில், ρ—மதிப்பிற்கேற்ப, $[1 + (n' - 1)\rho] = \theta$ (கருதியது) மதிப்பானது மாறுபடும்.

$\rho > 0$: n' அதிகரிக்கையில், θ மதிப்பானது அதிகரிக்கும். $\therefore$ V_c யின் மதிப்பானது மிகவும் அதிகமாகும். இந்நிலையில், n'-ஐ சிறியதாகவும், m-ஐப் பெரிதாகவும் கொள்வது பொருந்தும்.

$\rho < 0$: n' அதிகமாக இருப்பின், V_c-யானது சிறிய அளவில் அமைய n'-ஐ அதிக அளவிலும், m-ஐ சிறிதாகவும் கொள்வது நன்று.

ρ மதிப்பானது படிப்படியாக அதிகமாகி, $\rho = 1$ என்ற மதிப்புக்கு நெருங்குகையில், (m, n' இவற்றை நிலைப்படுத்தி),

திரள்களின் உள்ளீட்டுக் கூறுகள் ஓரினத் தன்மையுடன் விளங்க நேரிடுவதால், திரள்களின் இடையில் உள்ள வேறுபாடுகள் அதிகரிக்கின்றன இத்தகைய சூழ்நிலையில், V_c யானது அதிக மாகும். எனவே, $\rho > 0$ எனும்போது, n'-ஐக் குறைத்து, m-ஐ அதிகரிப்பது சிறந்ததாகக் கருதப்படுகிறது.

மாறாக, ρ-வின் மதிப்பு பூச்சியமாகவோ அல்லது அதற்கு நெருக்கமாகவோ இருப்பின், திரள்களில் உள்ள மாதிரிக் கூறுகள் தம்முள் அதிக அளவில் மாறுபட்டு விளங்கும். அதாவது, திரள்களின் உள்ளீட்டு மாறுபாடு (within variation) அதிகமாகிறது. இத்தகைய நிலையில், திரள்கள் அனைத்தும் ஒன்றுபோல் இருப்பதால், அவற்றின் இடைப்பட்ட மாறுபாடு (between variation) மிகவும் குறைகிறது. இந்நிலையில், சில திரள்களை மட்டுமே கருதி, அதிக அளவில் துணைக் கூறுகளை (subsamples) எடுத்தல் வேண்டும்

$\rho = 0$ எனில், n', m ஆகியவை V_c-யின் மதிப்பை எந்த விதத்திலும் பாதிப்பதில்லை.

$\rho < 0$ ஆனால், n' மதிப்பு அதிகரிக்கப்படுகையில் θ-வின் மதிப்பு குறைகிறது.

உதாரணமாக, $\rho = -0{\cdot}2$ எனில்,

$n' = 2$ எனக் கொண்டு, $\theta = 1 + (n'-1)\,\rho = 0{\cdot}8$

$n' = 5$ எனக் கொண்டு, $\theta = 1 + (5-1)\ (-0{\cdot}2) = 0{\cdot}2$ என்றவாறு அமைகிறது.

எனவே, V_c, V_{ran} இவற்றை ஒப்பிட, ρ-வின் மதிப்பானது பூச்சியமாக அல்லது அதற்கு நெருக்கமாய் இருப்பின், திட்ப வேறுபாடு (difference in precision) மிகக் குறைகிறது.

$\rho = 0$ எனில், திரள் மாதிரி எடுத்தல், சம வாய்ப்பு மாதிரி எடுத்தலைப்போன்றே இருக்கும். எனவே, மாதிரிக் கூறுகள் தம்முள் பலபடித்தானவைகளாக (Heterogeneous) இருப்பின், V_c-ன் மதிப்பீடாக, $\frac{S^2}{mn'}$ $\left(\text{அல்லது அதன் மாதிரி அளவை } \frac{s^2}{mn'}\right)$ என்பதைப் பயன்படுத்தலாம். எனவே ρ மதிப்பை மிகக் குறைவாகவும், திரள்களின் உள்ளீட்டு மாறுபாடு அதிகமாய் இருக்கும்படியும் அமைவது சாலச் சிறந்தது. வாழ்வு வசதி

பற்றிய ஆய்வுகளில் பெரிய நகரங்களைக் கருதி, அதிக அளவில் மாதிரிக் கூறுகள் (குடும்பங்கள்) தெரிவு செய்வது சிறந்த முடிவு களைத் தரும்.

ஆனல் பயிர் விளைவு பற்றிய ஆய்வுகளில், பரப்புகளை (m) அதிக அளவில் தெரிவு செய்து, துணைக் கூறுகளை (ஏகரில் காணப்படும் விளைச்சலைக் கருதலாம்) சிறிய அளவினதாகக் கொண்டால் (n), நல்ல முடிவுகள் கிடைக்கும்.

(viii) திரள் மாதிரி முறையில் மாதிரிப் பங்கீடு : (Allocation of Sample in Cluster Sampling)

சம வாய்ப்பு மாதிரி முறையோடு ஒப்பிடுகையில், திரள் முறையில் m, n' இவற்றைப் பங்கிடும் முறையை ஓரளவு குறிப்பிட்டோம். ஆனல், மாதிரி எடுத்தலில் ஆகும் செலவைப் பற்றி (cost) எதுவும் கருதப்படவில்லை. இங்கு, கொடுக்கப்படும் செலவுத் திட்டத்திற்கேற்ப, m, n' இவற்றை உத்தமமான அளவில் (optimum) எவ்வாறு பங்கிடுவது என்பது பற்றிய குறிப்பினைக் கருதுவோம் உத்தமப் பங்கீட்டு முறைப்படி, கொடுக்கப்பட்ட செலவு வரையறைக்கேற்ற வீதத்தில், குறைந்த (சிறும) அளவில் (minimum) மாறுபாட்டினைப் பெற்றிருக்கும் பங்கீட்டு முறையைக் கருதுகிறேம்.

முதலில், செலவுச் சார்பினை வரையறுக்கிறேம். முதற் கட்டமாகிய திரள்களின் மாதிரி எடுத்தலிலும், இரண்டாம் கட்டமாகிய எடுக்கப்படும் திரள் ஒவ்வொன்றிலிருந்தும் மாதிரி எடுத்தலிலும், தனித்தனியாக செலவுகள் அமைய நேரிடுவதால், இரண்டு வகைச் செலவினையும் இணைக்கும் வகையில், செலவுச் சார்பினை வரையறுத்தல் வேண்டும்.

c_1 என்பது முதல் கட்டத் தெரிவில் ஆகும் சராசரி செலவெனவும், c_2 என்பது இரண்டாம் கட்டத் தெரிவில் சராசரி யாக ஆகும் செலவெனவும் கொள்க. மேலும், செலவுச் சார்பினை நேர்கோட்டு அமைப்பில் (straight line) கருதுக. மாருத செலவினை c_0 என்று குறிப்பிட, மொத்தச் செலவுச் சார்பானது

$$c = c_0 + c_1.m + c_2.mn'$$

என்று தரப்படும். ஆனல், மாதிரிப் பங்கீடுபற்றிய பிரச்சினையில் c_0 மதிப்பு அதிகம் பாதிக்காததால், அதனை நீக்கிவிட்டு

$$c = c_1.m + c_2.mn'$$

என்றே கருதலாம். இனி உத்தமப் பங்கீடு பற்றிய பிரச்சினையை விரிவாக ஆராயலாம்.

திரள் மாதிரி முறை மாறுபாடானது

$$V(\hat{Y}) = \frac{M-m}{M} \cdot \frac{S_{1b}^2}{m} + \frac{N'-n'}{N'} \cdot \frac{S_{2i}^2}{mn'}$$

என்று ஏற்கெனவே பெறப்பட்டது. எனவே, உத்தமப் பங்கீட்டுப் பிரச்சினையானது, கொடுக்கப்பட்ட செலவுத் திட்டத்தில்(cost function), $V(\hat{Y})$-ன் மதிப்பைச் சிறுமஅளவின தாக்கத்தக்க (minimise) m, n' மதிப்புகளைக் கணக்கிடுவதாக உருப் பெறுகிறது. இதனை, இலக்ராஞ்ச் பெருக்கற் கெழுவான λ-வினைப் பயன்படுத்தி, பின்வருமாறு காணலாம் :

மாறுபாட்டு அளவீடான $V(\hat{Y})$-ஐயும், செலவுச் சார்பையும் இணைத்து,

$$F = V(\hat{Y}) + \lambda\,(c_1.m + c_2.mn' - c)$$

என்றும் சார்பினைக் கொள்வோம்.

இதிலிருந்து, வகையீட்டு விதிப்படி,

$$\frac{\partial F}{\partial m} = 0 \text{ என்பதிலிருந்து, } \left[-\frac{S_{1b}^2}{m^2} - \frac{S_{2i}^2}{m^2 \cdot n'} + \frac{S_{2i}^2}{N'.m^2} + \lambda\,(c_1 + c_2.n')\right] = 0 \quad \ldots (1)$$

எனவும்,

$$\frac{\partial F}{\partial n'} = 0 \text{ என்பதிலிருந்து, } \left[-\frac{S_2{}^2}{m\,(n')^2} + \lambda\; c_2.m\right] = 0 \quad \ldots (2)$$

எனவும் பெறுகிறோம்.

சமன்பாடுகள் (1), (2) ஆகியவற்றைத் தீர்ப்பதின் மூலமாக,

$$\lambda m^2 = \frac{S_1^{b2} - \left(\frac{S_{2i}^2}{N'}\right)}{c_1}$$

என்றும்,

$$\lambda m^2 = \frac{S_{2i}^2}{c_2 \cdot (n')^2}$$

என்றும் பெறலாம்.

இவ்விரண்டினையும் சமப்படுத்தி,

$$(n')^2 = \frac{c_1}{c_2} \cdot \frac{S_{2i}^2}{S_{1b}^2 - \left(\frac{S_{2i}^2}{N'}\right)}$$

$$\text{அல்லது } n' = \sqrt{\frac{c_1}{c_2} \cdot \frac{S_{2i}^2}{[S_{1b}^2 - (S_{2i}^2/N')]}} \quad \ldots \text{(A)}$$

என்று அறிகிறோம். பெறப்பட்ட n-ன் மதிப்பை செலவுச் சார்பில் பிரதியிட, m-ன் மதிப்பும் தீர்மானிக்கப்படும்.

n'-ஐ விட மிகவும் பெரிய அளவில் N'-ஐத் தெரிவு செய்யும் நிலையில் மாறுபாடானது

$$V(\hat{\bar{Y}}) = \frac{M-m}{M} \cdot \frac{S_{1b}^2}{m} + \frac{S_{2i}^2}{mn'}$$

என்ருகும். முன்போலவே, λ-வினைப் பயன்படுத்தி,

$$n' = \sqrt{\frac{c_1}{c_2} \cdot \frac{S_{2i}^2}{S_{1b}^2}} \quad \ldots\ldots \text{(B)}$$

என்று பெறலாம். Bயிலிருந்து $c_1 > c_2$ என்று அமையும் நிலையில், இரண்டாம் கட்டத்தேர்வு அதிக அளவில் அமையவேண்டும். அதாவது, திரள்களைத் தெரிவு செய்ய அதிகம் செலவாகும் பட்சத்தில், எடுக்கப்பட்ட திரள்களினின்று துணைக்கூறுகளை (subunits) அதிக அளவில் எடுத்தல் அவசியமாகிறது. மேலும் $S_{2i}^2 < S_{1b}^2$ என்று அமைந்திடின், துணைக்கூறுகள் குறைந்த அளவினவாதல் போதுமானது. அதாவது, திரள்கள் ஓரினத் தன்மையதாய் இருப்பின், இரண்டாம் கட்டத்தில் எடுக்கப்படும் துணைக்கூறுகள் குறைந்த எண்ணிக்கையாதல் (size) வேண்டும். முன்பே, திரள் உள்ளீட்டு ஒட்டுறவுக் கெழு அடிப்படையில் பெற்ற முடிவுக்கு இம் முடிவானது ஒத்த விதத்தில் அமைவது குறிப்பிடத்தக்கது.

முதலில்பெறப்பட்ட n'-ன்மதிப்பை அதாவது A-யினை, பிரிவு—உள்ளிட்ட ஒட்டுறவுக்கெழு, ρமூலமாகப் பின்வருமாறு பெறலாம்:

தேற்றம் :

கொடுக்கப்பட்ட செலவுத் திட்டத்தில், மாறுபாட்டை குறைக்கும் வகையில் அமையும் துணைக்கூறுகளின் அளவானது, (அதாவது, n')

$$n' = \sqrt{\frac{c_1}{c_2} \cdot \frac{1-\rho}{\rho}}$$

என அமையும்.

நிரூபணம் :

முன்பே கருதியது போல்,

$$S_{1b}^2 = \frac{(MN'-1)}{(N')^2 (M-1)} \cdot S^2 \cdot [1 + (N'-1) \cdot \rho]$$

$$S_{2i}^2 = \frac{(MN'-1)}{MN'} \cdot S^2 \cdot (1-\rho)$$

இவை இரண்டினையும், பின்வரும்விதத்தில் தோராய அளவுகளாக எழுதலாம் :

அதாவது,

$$S_{1b}^2 = \frac{MN'}{(N')^2 M} \cdot S^2 \cdot [1 + (N'-1) \cdot \rho]$$

$$= \frac{1}{N'} \cdot S^2 \cdot (1 + \overline{N'-1} \cdot \rho)$$

மற்றும் $S_{2i}^2 = S^2 \cdot (1-\rho)$

என்றபடி அமையும்.

எனவே, (A) — என்ற அளவீட்டில்,

$$S_{1b}^2 - \frac{S_{2i}^2}{N'} = \frac{1}{N'} \cdot S^2 (1 + \overline{N'-1} \cdot \rho) - \frac{1}{N'} \cdot S^2 (1-\rho)$$

$$= \frac{1}{N'} \cdot S^2 [1 + (N'-1) \rho - 1 + \rho]$$

$$= \frac{1}{N'} \cdot S^2 \cdot (N' \cdot \rho) = S^2 \cdot \rho.$$

என்ற மதிப்பினைப் பயன்படுத்த,

$$n' = \sqrt{\frac{c_1}{c_2} \cdot \frac{S^2 (1-\rho)}{S^2 \cdot \rho}}$$

$$= \sqrt{\frac{c_1}{c_2} \cdot \frac{1-\rho}{\rho}}$$

என்று பெறப்படும். இதன்மூலமாக, ρ-வின் மதிப்பு அதிகரிக்கையில், n'-ன் மதிப்பு குறைவதை அறியலாம். இது முன்பே காணப்பட்ட உண்மையாகும்.

எடுத்துக்காட்டு :

ஒரு திரளினைச் சென்று அடைவதற்கு ஆகும் செலவு ரூ.45 என்றும், அதில் உள்ள உறுப்பினர்களைக் கண்டு விவரங்கள் அறிய ஆகும் செலவு சராசரியாக ரூ.5 என்ற நிலையில், மாறாச் செலவு ரூ200 எனவும், மொத்த செலவுத்திட்டம் ரூ. 1200 எனவும் கொண்டு, m, n' இவற்றின் மதிப்பீடுகளைப் பெறுக. பழைய மதிப்பீடுகளினின்று, ρ—வானது 0·1 என்றும் அறியப்படுகிறது.

மேற்கண்ட உண்மைகளை, இந்நிலையில் பயன்படுத்தலாம்.

$C = 1200$;	$c_0 = 200$.
$c = 45$;	$c_2 = 5$.
$\rho = 0{\cdot}1$;	

செலவுச் சார்பின்படி,

$$C = c_0 + c_1 m + c_2 . mn'$$

அதாவது $1200 = 200 + 45m + 5mn'$ .... (1)

என்பதாகும்.

மேலும் இறுதியாகக் கண்ட முடிவின்படி,

$$n' = \sqrt{\frac{c_1}{c_2} \cdot \frac{1-\rho}{\rho}} \text{ என்பதாகும்.}$$

$$n' = \sqrt{\frac{45}{5} \cdot \frac{(1-0{\cdot}1)}{0{\cdot}1}}$$

$$= \sqrt{9 \times \frac{0{\cdot}9}{0{\cdot}1}} = \sqrt{9 \times 9} = 9 \text{ ஆகும்.}$$

சமன்பாடு I-லிருந்து, $n' = 9$ எனப் பிரதியிட்டு,

$$1200 = 200 + 45m + 5mn'$$
$$= 200 + 45m + 5m.\ (9)$$

அல்லது $90m = 1000$

$$\therefore\ m = \frac{1000}{90} = \frac{100}{9} \simeq 22$$

என்று முடிவெடுக்கலாம்.

II. திரள் முறையில் வீத மதிப்பீடு (Estimation of Proportion in cluster Sampling)

பலதுறைகளில், திரள் மாதிரிமுறையினைப் பயன்படுத்தி, வீத மதிப்பீடுகள் (estimation of proportion) கருதப்படுவதாக அமைந்திருக்கும். உதாரணமாக, மொத்தத் தொகுதியினில், ஏதேனும் ஒருநோயினுல் பாதிக்கப்பட்டோரின் அளவுவீதமானது மதிப்பிட வேண்டியிருக்கலாம். சில வட்டப் பரப்புகளைத் தெரிவு செய்து, அவற்றிலிருந்து இரு வேறு நிலைகளில் மாதிரிகள் எடுத்து ஆய்வு நடத்தப்படலாம். துணைக் கூறுகளினின்று பெறப்பட்ட முடிவுகளை வைத்து, முழுமைத் தொகுதிபற்றிய உண்மைகள் அறியப்படலாம்.

பொருள் விற்பனை பற்றிய ஆய்வு, ஒரு குறிப்பிட்ட பிரச்சினை பற்றி மக்கள் கருத்தறியப்பயன்படும் ஆய்வு, தொழிலகங்களிலிருந்து வெளிப்படும் உற்பத்திப் பொருள்களினிடையே குறைபாட்டு வீதத்தை (proportion of defectives) அறிய மேற்கொள்ளப்படும் ஆய்வு, மக்களது வாழ்க்கைத்தரம் அறியும் பொருட்டுக் கருதப்படும் ஆய்வு, என்று பல்வேறு வகையான ஆய்வுகளில், வீத மதிப்பீடு தேவைப்பட்ட ஒன்ருக அமைகிறது.

குறியீடு :

இரு பிரிவாக ஒரு தொகுதியினைக் கருதுகையில், Y_{ij} என்ற மாறியினை,

$Y_{ij} = 1$ எனில், கருதப்படும் ஒரு பிரிவில் ஒரு கூறு (unit) இருப்பதாகவும்,

$Y_{ij} = 0$ எனில், அப்பிரிவில், அக் கூருனது இல்லாததாகவும், இருவகையாகக் கொள்ளுவோம். மொத்தம் M திரள்களாக முழுமைத் தொகுதியைப் பிரித்து,

$$Y_i = \sum_{N_i} Y_{ij} = A_i$$

என்று i-ஆவது திரளில் மொத்த மதிப்பைக் கருதி, அதன் மூலமாக,

$$P_i = \frac{Ai}{N_i} = \frac{y_j}{N_i}$$

என்று i—ஆவது திரளின் வீத அளவினைக் குறிக்கிறோம். (i 1, 2......M)

மேலும் $Q_i = 1 - P_i$ என்றும் குறிப்பிடுகிறோம்.

N_i-யிலிருந்து, n_i அளவுள்ள ஒரு துணைமாதிரி **(subsample)** எடுக்கப்படின், அதனுடைய வீத அளவு

$$p_i = \frac{a_i}{n_t} = \frac{y_i}{n_i} \quad \text{ஆகும்.}$$

A_i என்பது i-வது துணை மாதிரியில் அல்லது துணைக்கூறில் உள்ள y_{ij}-யின் மொத்த மதிப்பாகும்.

இத்தகைய வீதங்களை மதிப்பிடுவது, சமவாய்ப்பு முறையின் அடிப்படையில் அமையக் காணலாம்.

தொகுதியின் வீத மதிப்பீடு :

P_i என்ற துணைக்கூறு வீதங்களைப் பயன்படுத்தி, P என்னும் தொகுதி வீதத்தின மதிப்பிடலைக் கருதுகிறோம்.

$$\hat{\bar{Y}} = \frac{M}{Nm} \sum_{m} \frac{N_i}{n_i} \sum_{n_i} y_{ij} \quad \{y_{ij} = 0 \text{ அல்லது } 1\}$$

என்பதை ஏற்கெனவே கண்டறிந்தோம்.

இதன்படி,

$$\hat{\bar{Y}} = \frac{M}{N \cdot m} \sum_{m} N_i p_i$$

என்பது P-யின் மதிப்பீடாகும்.

அதாவது,

$$\hat{p} = \frac{M}{Nm} \sum_{m} N_i p_i$$

என்பது வீதமதிப்பீடாகும்.

$N_i p_i$ என்பது, i-ஆவது திரளில், குறிப்பிட்ட இயல்பினைப் பெற்றுள்ள உறுப்புகளின் எண்ணிக்கை மதிப்பாகும்.

$\sum_{m} N_i p_i$ என்பது, m—திரள்களில் (மாதிரிக் கூறு), அந்த இயல்பினைப் பெற்றிருக்கும் உறுப்புக்களின் எண்ணிக்கை மதிப்பாகும். இதன் அடிப்படையில், $\frac{1}{m} \sum_{m} N_i p_i$ என்பது ஓவ்வொரு திரளிலும், சராசரியாகக் காணப்படும் வீத அளவு மதிப்பீடு. மொத்தம் M திரள்கள் தொகுதியில் அமைவதால்,

∴ இவற்றுக்கேற்ற மொத்த மதிப்பீடு $\frac{M}{m} \sum N_i P_i$ எனப் பெறப்பட்டு, இதனை $\frac{1}{N}$—ஆல் பெருக்க, $\hat{p}$ பெறப்படுகிறது.

$$\text{மேலும் } E(\hat{p}) = E_i\left(E_j \frac{M}{Nm} \sum_{m} N_i p_i \right)$$

$$= E_i \left(\frac{M}{Nm} \sum_{m} N_i P_i \right)$$

$$= \frac{M}{Nm} \cdot m \sum_{M} \frac{1}{M} (N_i P_i)$$

$$= \frac{\sum_{M} N_i P_i}{N} = P$$

என்று அறியப்படுவதன்மூலம், $\hat{P}$ ஆனது P-யின் பிறழ்ச்சியற்ற மதிப்பீடாகிறது.

துணைமுடிவு :—

$$N' = N_i = \frac{N}{M} \text{ என்ற நிலையில்,}$$

$$\hat{P} = \frac{M}{Nm} \sum_{m} N_i P_i$$

$$= \frac{1}{N'm} \sum_{m} N' p_i$$

$$\therefore \hat{P} = \frac{1}{m} \sum_{m} p_i \text{ ஆகும்.}$$

∴ இதற்கு, p_i மதிப்புகளின் சராசரியினைக் காணவேண்டும்.

$\hat{P}$—ன் மாறுபாடு (Sampling Variance of $\hat{P}$)

$V(\hat{P})$-ஐக் கணக்கிட, முதலில் $V(\hat{\bar{Y}})$-ஐக் கருதி (திரள் மாதிரி அடிப்படையில் அதனை P, Q இவற்றின்மூலம் எழுத வேண்டும். ஏற்கெனவே அறிந்தபடி,

$$V(\hat{\bar{Y}}) = \frac{1}{N^2}\left(M^2 \cdot \frac{M-m}{M} \cdot \frac{S_b^2}{m} + \frac{M}{m} \sum_{M} N_i^2 \frac{N_i - n_i}{N_i} \cdot \frac{S_i^2}{n_i} \right)$$

$$= \frac{1}{(N')^2} \frac{M-m}{M} \frac{S_b^2}{m} + \frac{M}{m} \cdot \frac{1}{N^2} \sum_{M} N_i^2 \frac{N_i - n_i}{N_i} \cdot \frac{S_i^2}{e_i}$$

திரள்களின் இடைப்பட்ட மாறுபாடான

$$S^2 = \frac{1}{M-1} \cdot \sum_{M} (y_i - \bar{Y})^2$$

என்பதை,

$$S_b^2 = \frac{1}{M-1} \sum_{M} (N_i p_i - N' P)^2$$

எனஎழுதலாம். $\{\because \bar{Y}_i = N_i P_i$ என்றும்

$$\bar{Y} = \frac{Y}{M} = \frac{N}{M} \sum_{M} \frac{Y_i}{N} = N'P \text{ என்றும்}$$

அமைந்திடும்.$\}$

இதுபோல்,

$$S_i^2 = \frac{1}{N_i - 1} \sum_{j:1}^{N_i} (Y_{ij} - \bar{Y}_i)^2$$

என்பதனை,

$$S_i^2 = \frac{1}{N-1} (N_i \; P_i \; Q_i) \text{ என எழுதலாம்.}$$

ஏனெனில்,

$$\sum_{N_i} (y_{ij} - \bar{Y})^2 = N_i \; P_i \; Q_i \; \{\text{சமவாய்ப்புமுறை அடிப்படையில்}\}$$

எனவே, $\hat{P}$ – யின் மாறுபாடு

$$V.(\hat{p}) = \frac{M-m}{M(M-1)(N')^2} \sum_{1}^{M} \frac{(N_i P_i - N'P)^2}{m}$$

$$+ \frac{M}{mN^2} \sum_{1}^{M} \frac{N_i^2 (N_i - n_i)}{N_i} \frac{P_i Q_i}{n_i}$$

என்றாகும்.

$V(\hat{P})$ யில் வலப்புற, முதல் உறுப்பானது m திரள்களை M திரள்களினின்று சமவாய்ப்பு முறைப்படி எடுப்பதில் உள்ள மாறுபாட்டைக் குறிக்கும். $M = m$ ஆனால், இது பூச்சியமாகிறது.

இதேபோல் N_i கூறுகள் உடைய i-ஆவது துணை உறுப்பிலிருந்து n_i அளவுள்ள மாதிரி தெரிவு செய்யப்படின், அதன்மூலம் உள்ளீட்டு மாறுபாடு ஏற்படுகிறது. இத்தகைய மாறுபாடுகளின் மொத்தத்தை வலப்புற இரண்டாம் உறுப்பு காட்டுகிறது.

$\therefore$ $N_i = n_i$ ஆனால், இதன் மதிப்பும் பூச்சியம் ஆகிறது.

$M = m$ எனில், $V(\hat{P})$ ஆனது படுகை மாதிரிமுறை மாறுபாட்டிற்குச் சமமாக அமைகிறது.

தரப்பிழை அளவு (standard error), $\sqrt{V(\hat{P})}$ என்று பெறப்படும்.

$V(\hat{P})$-ன் மாதிரி மதிப்பீடு :

P_i, P என்ற தொகுதி மதிப்புகளை அறிந்தால்தான் $V(\hat{P})$ யினைக் காண இயலும். எனவே தகுந்த முறையில் மாதிரி மதிப்பீடுகளைப் பெறுவது அவசியமாகிறது.

$V(\hat{Y})$-ன் மாதிரி அளவையினைப்போன்றே, $V(\hat{P})$-ன் மாதிரி அளவையும் பெறப்படுகிறது.

($y_{ij} = 0$ அல்லது 1 என்று கொள்ளப்படும்).

முதல் மதிப்பீட்டு அளவை :

$V(\hat{Y}_c)$-ன் பிறழ்ச்சியற்ற மதிப்பீடானது

$$\hat{V}(\hat{Y}_c) = \frac{1}{N^2}\left(M^2 . \frac{M-m}{M} . \frac{S_b^2}{m} + \frac{M}{m}\sum_{m} N_i^2 . \frac{N_i - n_i}{N_i} . \frac{s_i^2}{n_i}\right)$$

இதில்,

$$s_b^2 = \frac{1}{m-1}\sum_{m}(\hat{Y}_i - \hat{Y})^2$$

$$s_i^{\,2} = \frac{1}{n_i - 1} \sum_{n_i} (\hat{y}_{ij} - \bar{y}_i)^2$$

$$\hat{Y}_i = N_i p_i \text{ என்றும், } \hat{Y} = \frac{1}{m} \sum_{m} N_i p_i$$

என்றும் ஏற்கெனவே பெறப்பட்டன.

$$\therefore \quad s_b^{\,2} = \frac{1}{m-1} \sum_{m} \left(N_i p_i - \frac{1}{m} \sum_{m} N_i p_i \right)^2$$

$$s_i^{\,2} = \frac{1}{(n_i - 1)} \sum_{n_i} (y_{ij} - \bar{y}_i)^2 = \frac{1}{n_i - 1} . \; (n_i p_i q_i)$$

$$\left[\because \sum_{n_i} (y_{ij} - \bar{y}_i)^2 = n_i p_i q_i \right.$$

எனவே,

$$\hat{V}(\hat{P}) = \frac{1}{N^2} \left[M^2 . \frac{M-m}{M} . \frac{1}{m} . \frac{1}{m-1} \sum_{m} \left(N_i p_i - \frac{1}{m} \sum_{m} N_i p_i \right)^2 + \frac{M}{m} \sum_{m} N_i^{\,2} . \frac{N_i - n_i}{N_i} . \frac{1}{n_i} . \frac{1}{n_i - 1} (n_i p_i q_i) \right]$$

ஆகும்.

$$\therefore \quad \text{தரப்பிழை : } s(\hat{P}) = \sqrt{\hat{V}(\hat{P})}.$$

இரண்டாம் வகை மதிப்பீடு :

$$N' = \frac{N}{M} = N_i \text{ என்றும்}$$

$$n_i = n' = \frac{n}{m}$$

என்றும் அமையும் நிலையாகும்.

$N_i = N'$ என்றும் ; $n_i = n'$ என்றும் பிரதியிட்டு,

$$\hat{V}(\hat{P}) = \frac{M-m}{M} \cdot \frac{1}{m} \cdot \frac{1}{m-1} \sum_{m} (p_i - \bar{p})^2 + \frac{N'-n'}{N'} \cdot \frac{1}{M} \cdot \frac{1}{m(n'-1)} \sum_{m} p_i q_i$$

என்று அமையும்.

இதில்,

$$\left[\hat{P} = \frac{M}{Nm} \cdot \sum_{m} N_i p_i = \frac{1}{m} \Sigma p_i = \bar{p} \text{ (அதாவது } p\text{-யின் சராசரி)}\right]$$

மேலும்,

$$\bar{p} = \frac{1}{mn'} \cdot \sum_{m} y_i.$$

மூன்றுவது வகை மதிப்பீடு :

$N' = N_i$; $n' = n_i$ என்பதோடு, m-உடன் ஒப்பிடுகையில், M அளவானது மிகப் பெரிதாக அமைய நேரிடின்,

$$\hat{V}(\bar{P}) = \frac{M-m}{M} \cdot \frac{1}{m} \cdot \frac{1}{(m-1)} \cdot \sum_{m} (p_i - \bar{p})^2 + \frac{N'-n'}{N'} \cdot \frac{1}{M} \cdot \frac{1}{m(n'-1)} \sum_{m} p_i q_i$$

$$\therefore \hat{V}(\hat{P}) \simeq \frac{1}{m} \cdot \frac{1}{(m-1)} \sum_{m} (p_i - \bar{p})^2$$

என்று சுருங்குகிறது.

$$\left\{ \because \frac{1}{M} \simeq 0 ; \quad \frac{M-m}{M} \simeq 1 \right\}$$

இவ்வாறு, மூன்று வகையான நிலைகளில் $\hat{V}$ $(\hat{P})$-ன் மதிப்புப் பெறப்படுவதைக் காணலாம்.

எடுத்துக்காட்டு 1

$M = 4$ பிரிவுகளாய் அமைந்த மாணவர்களுள், கல்லூரி செல்ல எண்ணுபவர்கள் வீதத்தை மதிப்பிட வேண்டும்.

பிரிவு அல்லது பகுதி	N_i	Y_i	P_i	y_i	$P_i = \frac{y_i}{n_i}$
1	5	3	$\frac{3}{5}$	3	$\frac{3}{4}$
2	4	3	$\frac{3}{4}$	0	0
3	7	4	$\frac{4}{7}$	2	$\frac{2}{4}$
4	12	6	$\frac{6}{12}$	3	$\frac{3}{4}$

(i) P-யினைக் கணக்கிடுதல் :

வரையறைப்படி : $P = \frac{Y}{N}$

அதாவது $$\frac{\sum_{M} Y_i}{N} = \frac{3+3+4+6}{28} = \frac{16}{28} = \frac{4}{7}$$

$(M = 4)$

$\therefore \quad P = \frac{4}{7}$ ஆகும்.

(ii) 1, 3, 4 எனும் பிரிவுகள் தெரிந்தெடுக்கப்பட்டு, ஒவ்வொரு பிரிவிலும் $n' = 4$ என்ற அளவில் மாணவர்கள் தேர்ந்தெடுக்கப்படின், இத் துணை மாதிரிகளின் மதிப்புகளை (y_i) பயன்படுத்தி, $\hat{P}$-ஐக் கணக்கிடல்.

$m = 3$; $n' = 4$.

$$\therefore \quad \hat{P} = \frac{M}{Nm} \sum_{m} N_i p_i$$

$$= \frac{4}{28 \times 3}\left[\left(5 \times \frac{3}{4}\right) + \left(7 \times \frac{2}{4}\right) + \left(12 \times \frac{3}{4}\right)\right]$$

$$= \frac{1}{21}\left[\frac{15}{4} + \frac{14}{4} + \frac{36}{4}\right]$$

$$\therefore \quad \hat{P} = \frac{1}{21} \cdot \frac{65}{4} = \frac{65}{84}.$$

(iii) மேற்கண்ட $\hat{P}$-ன் மாறுபாடு காணல் :

அதாவது $V(\hat{P})$-ஐக் கணக்கிடல் :

$$V(\hat{P}) = \frac{M - m}{M(M-1)\,.\,(N')^2} \sum_{1}^{M} \frac{(N_i P_i - N' P)^2}{m}$$

$$+ \frac{M}{mN^2} \sum_{1}^{M} \frac{N_i^2 (N_i - n_i)}{(N_i - 1)} \times \frac{P_i Q_i}{n_i}$$

$$N' = \frac{N}{M} = \frac{28}{4} = 7.$$

$P = \frac{4}{7}$ என ஏற்கெனவே அறியப்பட்டது.

$\therefore \quad N' P = \left(\frac{4}{7} \times 7\right) = 4$ ஆகும்.

(திரள்) பிரிவு	N_i	P_i	$N_i P_i$	$P_i Q_i$	$(5)/n'$	$\frac{N_i^2(N_i - n_i)}{(N_i - 1)}$	$(6)\times(7)$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	
1	5	$\frac{3}{5}$	3	$\frac{6}{15}$	$\frac{1}{10}$	$\frac{25}{4}$	$\frac{5}{8}$
2	4	$\frac{3}{4}$	3	$\frac{3}{16}$	$\frac{3}{64}$	0	0
3	7	$\frac{4}{7}$	4	$\frac{12}{49}$	$\frac{3}{49}$	$\frac{49}{2}$	$\frac{3}{2}$
4	12	$\frac{6}{12}$	6	$\frac{36}{144}$	$\frac{1}{16}$	$\frac{1152}{11}$	$\frac{72}{11}$
							8·67

இங்கு $n_i = n' = 4$ ஆகிறது.

$m = 3$.

$$\therefore \quad V(\hat{P}) = \frac{(4-3)}{4 \times 3 \times (7)^2} \cdot \frac{1}{3} [(3-4)^2 + (3-4)^2$$

$$+ (4-4)^2 + (6-4)^2] + \frac{4}{3 \times (28)^2} \cdot (8{\cdot}67)$$

$$= \left[\frac{1}{1764} \times 6\right] + \frac{8{\cdot}67}{588}$$

$$\therefore \quad V(\hat{P}) = 0{\cdot}003 + 0{\cdot}015$$

$$\therefore \quad V(\hat{P}) = 0{\cdot}018.$$

என மதிப்பிடப்படுகிறது.

எடுத்துக்காட்டு 2

$M = 15$; $m = 6$; $N = 150$; $n' = 4$.

திரள் எண்	N_i	y_{ij}	$p_i = \frac{\bar{y}_i}{n_i}$	$N_i\, p_i$
1	14	2	$\frac{1}{2}$	7
2	12	3	$\frac{3}{4}$	9
3	24	3	$\frac{3}{4}$	18
4	6	1	$\frac{1}{4}$	1·5
5	8	4	1	8
6	13	1	$\frac{1}{4}$	3·25
				46·75

$\hat{P}$-யின் மதிப்பீடு :

$$\therefore \quad \hat{P} = \frac{M}{Nm} \sum_{i}^{m} N_i\, p_i$$

$$\therefore \quad \hat{P} = \frac{15}{(150 \times 6)} \cdot (46{\cdot}75) = 0{\cdot}78.$$

பயிற்சி : இம் மாதிரியினைக் கருதி, $V(\hat{P})$-ன் மதிப்பீட்டினை (மாதிரிகளின் அடிப்படையில்) காண்க. அதாவது, $\hat{V}(\hat{P})$ காண்க.

குறிப்பு : $V(\hat{P})$-யில் P_i, Q_i-க்குப் பதில், p_i, q_i இவற்றைப் பயன்படுத்தவும்.

எடுத்துக்காட்டு : 3

500 பள்ளிகளிலிருந்து, உயர்நிலைக்கல்வி பயிலும் மாணவர்களிடை, அடிக்கடி படக்காட்சிக்குச் செல்லும் பழக்கம் உடையோர் விகிதத்தினை (proportion) மதிப்பிட, $m = 10$ பள்ளிகள் தெரிந்தெடுக்கப்பட்டு, ஒவ்வொரு பள்ளியிலிருந்தும் 20 மாணவர்கள் தேர்ந்தெடுக்கப்படுகின்றனர். ஒழுங்கு மாதிரி முறை மேற்கொள்ளப்படுவதாகக் கொள்வோம். அதன்படி, $\frac{M}{m} = \frac{500}{10}$ $= 50$ என்ற மதிப்பு கிடைக்கும். 1 முதல் 50 க்குள் ஒரு ராண்டம் எண்ணைத் தெரிவுசெய்வோம். உதாரணமாக, '60' என்ற எண்ணை எடுப்பதாகக் கொள்வோம்.

ஒழுங்கு மாதிரியானது (systematic sample),

60, 110, 160, 210, 260, 310, 360, 410, 460, 510 என்ற எண்கள் கொண்ட பள்ளிகள் தெரிவில் இடம் பெறும். இவற்றிலிருந்து ஒவ்வொன்றிலிருந்தும் 20 ($n' = 20$) மாணவர்களைக் கருதிப் பெறப்படும் மதிப்புகள் பின்வருமாறு கொடுக்கப்படுவதாகக் கருதுவோம் :

m	y_i	p_i	$(p_i - \bar{P})$	$(p_i - \bar{P})^2$
1	7	0·35	0·09	0·0081
2	4	0·20	−0·06	0·0036
3	5	0·25	−0·01	0·0001
4	4	0·20	−0·06	0·0036
5	7	0·35	0·09	0·0081
6	5	0·25	−0·01	0·0001
7	3	0·15	−0·11	0·0121
8	6	0·3	0·04	0·0016
9	6	0·3	0·04	0·0016
10	5	0·25	−0·01	0·0001
$n' = 20$ (எல்லாத் திரள்களுக்கும்)				0·0390

$P_i = \frac{y_i}{n_i}$ (ஒவ்வொரு திரளிலும்).)

$$\therefore \quad \bar{p} = \frac{\Sigma p_i}{10} = \frac{2 \cdot 60}{10} = 0 \cdot 26$$

$$\hat{P} = \frac{1}{mn'} \sum_m y_i = \frac{52}{200} = 0 \cdot 26$$

மேலும்,

$$\hat{V}(\hat{p}) = \frac{1}{m(m-1)} \sum_m (p_i - \bar{p})^2$$

$$= \frac{1}{10 \times 9} (0 \cdot 039)$$

$$= 0 \cdot 0004$$

இதிலிருந்து,

$$s(\hat{p}) = \sqrt{0 \cdot 00\,04} = 0 \cdot 02.$$

பயிற்சி

1. இரவு 8 மணிக்கு மேலும் அலுவல் நடைபெறும் விற்பனை யகங்களின் விகிதத்தை அல்லது வீதத்தை மதிப்பிட, சில தெரிந் தெடுக்கப்பட்ட தெருக்கள் பற்றிய விவரம் பின்வருமாறு:

$y_{ij} = 1$ எனில், j என்ற கடையானது, i-ஆவது திரளினைச் சார்ந்து, 8 மணிக்கு மேல் திறந்திருப்பதாக கொள்க. இல்லையேல் $y_{ij} = 0$ ஆகிறது.

தெரு (திரள்) எண்	N_i	y_{ij}	y_i
5	14	1,1,0,1	3
8	12	0,1,0,1	2
10	24	0,0,1,0	1
11	6	1,1,0,1	3
12	8	1,1,1,1	4
15	13	0,0,1,1	2

(i) பல்வேறு p_i மதிப்புகளைக் கண்டு, $\hat{P}$-யினைக் கணக்கிடுக.

(ii) $\hat{\hat{P}} = \frac{M}{Nmn'} \sum_{m} N_i y_i$ என்பதின்மூலம் $\hat{P}$ ஐக் காண்க.

(iii) $V(\hat{P})$-ன் மாதிரி அளவையினைக் கணக்கிடுக.

2. $M = 3$ என்ற திரள்களில், கருதப்படும் வீடுகளில் வானொலிப் பெட்டி உண்டா, இல்லையா என்று ஆராயப்படுகிறது.

முடிவுகள் :

திரள்	N_i	Y_i
1	3	2
2	4	2
3	5	3

(i) P_i, Q_i இவற்றை கணக்கிட்டு, $V(\hat{P})$ மதிப்பை அறிக. ($n_i = n' = 2$ எனக் கொள்க.)

(ii) இதிலிருந்து, $m = 2$; $n' = 2$ என்றபடி ஏதேனும் மாதிரிகளைக் கருதி, $V(\hat{P})$ யின் மதிப்பீட்டு அளவையினைத் தீர்மானிக்கவும்.

3. $n_i = n'$ என அமைந்திடின்,

(i) $s_b^2 = \frac{1}{m-1} \cdot \frac{1}{(n')^2} \sum_{m} \left(N_i y_i - \frac{1}{m} \sum_{m} N_i y_i \right)^2$

(ii) $s_i^2 = \frac{y_i (n' - y_i)}{n'(n'-1)}$

என்று நிறுவுவதன்மூலம், $\hat{V}(\hat{P})$ யினைக் கொடுக்கவும்.

(iii) $\sum \left(N_i y_i - \frac{1}{m} \sum_{m} N_i y_i \right)^2$

$$= \Sigma (N_i y)^2 - \frac{1}{m} (\Sigma N_i y_i)^2$$

என்றும் காட்டுக.

4. ஒரு நிறுவனமானது, தினசரி பெறும் ஆர்டர்களில் ரூ. 500 க்கு மேல் உள்ள ஆர்டர்களின் வீதத்தை மதிப்பிட்டு அறிய எண்ணுகிறது. தினசரி $M = 200$ பைகள் என்றபடி அஞ்சல் வழி ஆர்டர்கள் பெறப்படுகின்றன. இவற்றுள் $m = 10$ பைகள் மாதிரியாக எடுக்கப்பட்டு, ஒவ்வொரு பையிலிருந்தும் $n' = 50$ ஆர்டர்கள் தெரிவு செய்யப்படுகின்றன. முடிவு பின்வரும் பட்டியலாக அமைகிறது :

பைகள்	N_i	y_i
1	220	5
2	210	6
3	230	4
4	180	5
5	190	6
6	200	3
7	240	5
8	170	5
9	190	4
10	190	4

தொகுதியானது மொத்தம் $N = 40,000$ ஆர்டர்களைக் கொண்டதாகக் கொடுக்கப்படின்,

(i) $\hat{P}$-மதிப்பைக் கணக்கிடுக.

(ii) $\hat{V}(\hat{p})$ மற்றும் $s(\hat{P})$ ஆகியவற்றை மதிப்பிடுக.

(iii) இதன்மூலம், P-க்கு 2σ நம்பக எல்லையைத் தீர்மானம் செய்க.

5. மேற்கண்ட பயிற்சி (4)-ல், $N_i = N' = 200$, என்று கருதி, $\hat{P}$, $\hat{V}(\hat{P})$ இவற்றைக் காண்க.

III. (பலகட்ட மாதிரிமுறை)
(Multi stage Sampling Method)

ஏற்கெனவே, திரள்மாதிரி முறையில், இரு கட்டங்களில் மாதிரிகள் எடுக்கப்படும் நிலையினைப்பற்றி விரிவாகக் குறிப்

பிட்டோம். முதல்நிலைக் கூறுகளை (primary units) தெரிவு செய்து, இறுதியில் இரண்டாம் நிலை கூறுகளை (secondary units) தெரிவு செய்து, மதிப்பீடுகள் பெறும் நிலையைக் கண்டோம். இதனை மேலும் விரிவுபடுத்தி, பொதுவான முறையினை 'பல கட்ட மாதிரி முறை' (multi stage sampling) எனக் குறிக்கலாம். ஒவ்வொரு கட்டத்திலும் அல்லது நிலையிலும் (Stage), முன் நிலையில் தெரிவு செய்யப்பட்ட திரள்களினின்று மாதிரி எடுக்கப் படுவதாகக் கருதலாம். செலவைப் பொறுத்த அளவில், இது திரள்களின்று மாதிரிக் கூறுகளை எடுக்கும் முறையினைவிட குறைந்த அளவு திறனுடனும், மாதிரிசார்ந்த மாறுபாட்டினைப் பொறுத்தமட்டில், திரள்களில் தெரிவு செய்யும் முறையைவிட அதிகத் திறனுடனும், இத்தகைய முறை விளங்குகிறது.

இத்தகைய பலகட்ட மாதிரித் தேர்வு முறையானது, பல வகையான நடைமுறைச் சூழ்நிலைகட்கு பொருத்தமாக அமையும் ஒரே வழிமுறையாக விளங்குகிறது. இறிதியாகக் கண்டறிய வேண்டிய கூறுகளைப்பற்றிய மாதிரி சட்ட' அமைப்பு (sampling frame) சரிவர, திருப்திதரத்தக்க விதத்தில் அமைக்கப் படாத நிலையில், அல்லது அத்தகைய கூறுகளுடன் தொடர்பு கொள்ள ஆகும் செலவு அதிகமாக இருக்கும் நிலையில்,இத்தகைய வழி முறை ஒன்றே எளிதாகவும், பின்பற்றத் தக்கதாகவும் இருக்கு ம்சமூக மற்றும் பொருளாதார நிலைகள்பற்றி நிகழ்த்தப் படும். ஆய்வுகளில்,இத்தகைய வழிமுறையேபெரிதும் பின்பற்றப் படுவதைக் காணலாம். இத்தகைய ஆய்வினில், குடும்பங்கள் இறுதிக் கூறுகளாக விளங்கும். ஒரு இடத்தல் வசிக்கும் குடும்பங்கள் பற்றிய விரிவான பட்டியலானது கிடைக்காமல் இருக்கலாம். அல்லது அப்பட்டியலானது தெளிவாக இல்லா திருக்கலாம். மாறாக, கிராமங்கள் பற்றிய விரிவான தொகுப்போ, அல்லது நகர்ப்புறம் பிரிவுகள் பற்றிய விரிவான பட்டியலோ கிடைக்கப் பெறலாம். அந் நிலையில், முதலாவதாக கிராமங்கள் மற்றும் நகர்ப்பிரிவுகள் கொண்ட மாதிரியினைத் தெரிவு செய்ய லாம். ஒவ்வொரு, தெரிவுசெய்யப்பட்ட, கிராமப்பகுதி மற்றும் நகர்ப் பிரிவுகள் கொண்ட மாதிரியினைத் தெரிவுசெய்யலாம். ஒவ்வொரு, தெரிவு செய்யப்பட்ட, கிராமப்பகுதி மற்றும் நகர்ப் புறம் இவற்றினின்று, குடும்பங்கள் அமைந்த மாதிரியினைத் தேர்ந் தெடுக்கலாம். இது இரண்டாம் நிலை. மாறாக, கிராமங்கள் குறித்த பட்டியல் அமையாது, சில கிராமங்கள் கொண்ட வட்டாரப் பரப்புகள் பற்றிய முழு விவரம் மட்டுமே கிடைப்ப தாகக் கொண்டால், மேலும் ஒரு நிலையில் மாதிரி தேர்ந்தெடுக்கப் படவேண்டியதாகிறது. இது மூன்றுகட்ட மாதிரித் தேர்வு முறை (three stage sampling) ஆகும்.

பொதுவாக, தனித்தனிக் கூறுகளைவிட, அவற்றின் தொகுப்புகள் (அல்லது திரள்கள்) பற்றிய விவரங்களே அதிக அளவில் கிடைக்கின்றன. இறுதிநிலையில் கூறுகளை மிக்க கவனத்துடன் எடுக்க இயலுகிறது. ஒவ்வொரு நிலையிலும் பெறப்பட்ட விவரங்கள் நல்ல முறையில் பயன்படுத்தப்படுகின்றன. பல்வேறு வகையான ஆய்வுகளில், இம் முறையானது பெரிதும் பயன்படுத்தப்படுகின்றது

இருகட்ட மாதிரி முறை (Two Stage Sampling Method)

இதில் ஏற்கெனவே கண்டறிந்த முடிவுகளைத் தொகுத்து பின்வருமாறு கொடுக்கிறோம். (பார்க்க : திரள் மாதிரி முறை).

முழுமைத் தொகுதிக் கூறுகள் M திரள்களாகப் (clusters) பிரிக்கப்பட்டு, i-ஆவது திரளினில் N_i கூறுகள் இருப்பதாகக் கொண்டு, இந்த M திரள்களை முதல்நிலை கூறுகளாகவும் திரட்கூறுகளினை 2-வது நிலைக்கூறுகளாகவும் கருதி, m திரள்களை மாதிரியாக (கொடுக்கப்படும் வாய்ப்பின் அடிப்படையில்) தேர்ந்தெடுத்து, இவ்வாறு தெரிவு செய்யப்படும் i-ஆவது திரளிலிருந்து, இரண்டாம் கட்டத்தில், n_i துணைக் கூறுகளை அல்லது வழிக் கூறுகளை, சில வாய்ப்புகளைக் கருதித் தேர்ந்தெடுக்கிறோம்.

y_{ij} என்பது i என்ற முதல்நிலைக் கூறிலிருந்து வரும் j என்ற வழிக்கூறின் (secondary unit) மதிப்பு என்று (j : 1, 2 ... n_i ; i : 1, ... m) கருதிடின், $\hat{y}_i = \sum\limits_{n_i} a_{ij}\, y_{ij}$

என்று, i-ஆவது திரளின் மொத்தம் மதிப்பிடப்படும்.

a_{ij} என்பது பெருக்கு காரணி (inflation factor) எனப்படும், (2-வது நிலையில்).

∴ முழுமைத்தொகுதி மொத்தத்தின் (total) மதிப்பீடு

$$\hat{Y} = \sum_{m} a_i \sum_{n_i} a_{ij}\, y_{ij}$$

(a_i என்பது முதல் கட்டத்தில் பெருக்கு காரணி.)

சம அளவில் வாய்ப்புகள் அமைகையில், Y-ன் பிறழ்ச்சியற்ற மதிப்பீடு

$$\hat{\hat{Y}} = \frac{M}{m} \sum_{m} \frac{N_i}{n_i} \sum_{j} y_{ij} \qquad \ldots (1)$$

என்பதாகக் கண்டோம். (திரள் முறை)

இதேபோல், முதல் கட்டத்தில் கூறுகள் p_i என்ற வாய்ப்புகளுடன் (i : 1, 2 ... M), திரும்பவைக்கும் முறையில் எடுக்கப்பட்டு, n_i வழிக் கூறுகள் P_{ij} என்ற வாய்ப்புகளுடன் எடுக்கப்படின் (j : 1, 2 ...... N_i),

$$\hat{Y} = \frac{1}{m} \sum_{m} \frac{1}{p_i} \cdot \frac{1}{n_i} \sum_{m_i} \frac{y_{ij}}{p_{ij}} \qquad \ldots \quad (2)$$

(P_{ij}, p_i என்பவை, மாதிரித் தேர்வில் உள்ள வாய்ப்புகள்).

பொதுவாக :

இதே முறையினை விரிவுபடுத்தி, பல கட்ட மாதிரித் தேர்வில் படிப்படியாகக் கருதப்படும் மதிப்பீட்டினைப் பெறலாம். பொதுவாக, r-கட்டத் தேர்வு நிலையில்,

$$\hat{Y} = \sum_{i_1 = 1}^{m} a_{i_1} \sum_{i_2 = 1}^{n_{i_1}} a_{i_1 i_2} \ldots \sum_{i_r = 1}^{n_{i_1} - i_{r-1}} a_{i_1 i_2 \ldots i_r} y_{i_1 i_2 \ldots i_r}$$

ஒவ்வொரு கட்டத்திலும் மாதிரித் தேர்வானது வாய்ப்பு அடிப்படையில் அமைவதால், பல கட்டத் தேர்வில், வெவ்வேறு கட்டங்களில் தெரிவு முறைகளை நன்கு கருத்தில் கொண்டு, மதிப்பீடுகளின் எதிர்பார்க்கும் மதிப்பு மற்றும் மாறுபாட்டு அளவு ஆகியவற்றைக் காணுதல் வேண்டும். இறுதிக் கட்டத்தில் கருதப்படும் கட்டுப்பாட்டு எதிர்பார்க்கப்படும் மதிப்பானது (conditional expectation) அதற்கு முன் நிலையில் எடுக்கப்பட்ட கூறுகளைக் கொண்டும், அவற்றின் எதிர்பார்க்கும் மதிப்புகளை, முந்தைய நிலையில் எடுக்கப்படும் கூறுகளைக் கருதியும்......இதே மாதிரியாக மதிப்புகள் கணக்கிடப்படும். இறுதி நிலையில் தொடங்கி, பின்நோக்கிச் சென்று, கடைசியாக முதல் நிலையினை அணுகுதல் வழிமுறையாகும்.

இரு கட்டமாதிரி முறையில்,

$E_{12}(\hat{Y}) = E_1 E_2(\hat{Y})$ என்று சராசரி மதிப்பீடும்,

$V_{12}(\hat{Y}) = V_1 E_2(\hat{Y}) + E_1 V_2(\hat{Y})$ என்று மாறுபாட்டு மதிப்பும் கணக்கிடப்படும். 1, 2 என்ற துணைக்குறிகள் முதல், இரண்டாம் கட்டங்களைக் குறிக்கும்.

மூன்று கட்ட மாதிரித் தேர்வில் (Three stage Sampling),

$$E(\hat{Y}) = E_1 E_2 E_3 (\hat{Y})$$

என்றும்

$$V(\hat{Y}) = V_1.E_2 E_3(\hat{Y}) + E_1.V_2.E_2 E_3(\hat{Y}) + E_1 E_2.V_3(\hat{Y})$$

என்றும் காணப்படும். [மூன்று உறுப்புகளைக் காண்க.]

{ $V_1 E_2 (\hat{Y})$ என்பது முதல் நிலைக் கூறுகளின் மாறுபாடு ;

$E_1 V_2 (\hat{Y})$ என்பது முதற்கூறுகளில் அமையும் வழிக் கூறுகளின் இடைப்பட்ட மாடுபாடு — என்று பொருள் கொள்ள வேண்டும். }

பொதுவாக, r-கட்டங்கள் கொண்ட மாதிரித் தேர்வில் மொத்த மாறுபாடு r-உறுப்புகள் கொண்டதாக அமைந் திருக்கும். அதாவது ஒவ்வொரு கட்டத்திலும் அமையும் கூறுகளின் இடைப்பட்ட மாறுபாட்டினைத் (முன் நிலையில் கருதப் பட்ட கூறுகளினுள்) தரும் வண்ணம் ஒவ்வொரு உறுப்பும் அமைந்திருக்கும்.

எளிய ராண்டம் முறையில் (ஈடு செய்யப்படாத நிலையில்) இரு கட்ட மாதிரி முறை : (மாறுபடும் திரள் அளவு)

முன்பே அறிந்தபடி,

$$\hat{Y} = \frac{M}{m} \sum_{m} \frac{N_i}{n_i} \sum_{n_i} y_{ij}$$

$[E(\hat{Y}) = Y]$ M திரள்கள், ஒவ்வொன்றும் N_i அளவுள்ளவை.

$$\therefore \quad \bar{\bar{Y}} = \frac{1}{M \cdot N_0} \cdot \sum_{1}^{M} \sum_{1}^{N_i} Y_{ij} = \frac{1}{MN_0} \sum_{1}^{M} N_i \cdot \bar{Y}_i$$

$$\left(N_0 = \Sigma \frac{N_i}{M}\right)$$

என்று மதிப்பிடப்படும்.

ஈடு செய்யப்படாத முறையை (Without replacement) இரு கட்டங்களிலும் பயன்படுத்தி, மாறுபாட்டு மதிப்பினை

$$V(\hat{Y}) = V_1\, E_2(\hat{Y}) + E_1\, V_2(\hat{Y})$$

என்ற விதிப்படி ['1' என்பது முதல் கட்டத் திரள்களை எடுப்பதையும், '2' என்பது இரண்டாம் கட்ட துணை மாதிரிகள் எடுப்பதனையும் குறிக்கும்.] பின்வருமாறு பெறுகிறோம் :

$$\therefore\ V(\hat{Y}) = V_1\left[\frac{M}{m}\sum_1^m y_i\right] + E_1\left[\frac{M^2}{m^2}\sum_1^m N_i^2(1-f_i)\cdot\frac{S_i^2}{n_i}\right]$$

$$= M^2.(N_0)(1-f)\cdot\frac{S_b'^2}{m} + \frac{M}{m}\sum_1^M N_i^2(1-f_i)\cdot\frac{S_i^2}{n_i}$$

என்பதாகும்.

இதில்

$$S_b'^2 = \frac{1}{M-1}\sum_1^N\left(\frac{N_i}{N_0}\cdot\bar{\bar{Y}}_i - \bar{Y}\right)^2;$$

$$S_i^2 = \frac{1}{N_i-1}\sum_1^{N_i}(Y_{ij} - \bar{\bar{Y}}_i)^2;$$

$$f = \frac{m}{M};\quad f_i = \frac{n_i}{N_i};\quad N_0 = \frac{1}{M}\sum_1^M N_i$$

அதாவது N_0 என்பது ஒவ்வொரு முதற்கூறிலும் சராசரியாக அமையும் வழிக் கூறுகள் (secondary units). ($\because$ N_i மாறக் கூடியதாகக் கருதப்படும்.)

$S_b'^2$ என்பது திரள்களின் (முதல் நிலை) இடைப்பட்ட மாறுபாடாகவும், $S'_i{}^2$ என்பது திரளின் உள்ளீட்டு (within) மாறுபாடாகவும் வரையறுக்கப்படும். இவற்றை, இரண்டினையும் பிறழ்ச்சியற்ற முறையில் மதிப்பிட முடிந்தால், $V(\hat{Y})$-ன் பிறழ்ச்சியற்ற மதிப்பீட்டைப் பெறலாம்.

S_i^2-ன் பிறழ்ச்சியற்ற மதிப்பீடானது எளிதாக

$$s_i^2 = \frac{1}{(n_i - 1)} \sum_{1}^{n_i} (y_{ij} - \bar{y}_i)^2$$

என்று பெறப்படுகிறது.

மேலும், $S_b'^2$-ன் பிறழ்ச்சியற்ற மதிப்பீடு, $\sum_{1}^{M} N_i^2 \cdot \bar{Y}_i^2$ மற்றும் Y^2 இவற்றைப் பிறழ்ச்சியற்று மதிப்பிடுவதைப் பொறுத்ததாகும்.

$$S_b'^2 = \frac{1}{(M-1)\cdot(N_0)^2} \cdot \left[\sum_{1}^{M} N_i^2 \cdot \bar{Y}_i^2 - \frac{Y^2}{N} \right]$$

என்று எழுதி, பின்வருமாறு கருதலாம் :

நியதி அடிப்படையில்,

$V_2(\bar{y}_i) = E_2(\bar{y}_i^2) - \bar{Y}_i^2$ என்று பெறுவதால், $\bar{Y}_i^2$-ன் பிறழ்ச்சியற்ற மதிப்பீடு,

$$\bar{y}_i^2 - v_2(\bar{y}_i) = \bar{y}_i^2 - (1 - f_i) \cdot \frac{s_i^2}{n_i}$$

என்று வருவதால், இதிலிருந்து $\sum_{1}^{M} N_i^2 \cdot \bar{Y}_i^2$-ன் பிறழ்ச்சி அற்ற மதிப்பீடு

$$\frac{M}{m} \sum_{1}^{m} N_i^2 \left[\bar{y}_i^2 - (1 - f_i) \cdot \frac{s_i^2}{n_i} \right]$$

என்றாகும்.

Y^2-ன் பிறழ்ச்சியற்ற மதிப்பீடு, $\hat{Y}^2 - v\ (\hat{Y})$ ஆகும்.

இவற்றின் மூலம், $S_b'^2$-ன் பிறழ்ச்சியற்ற மதிப்பீடு

$$\frac{M}{m\,(M-1)} \cdot \left[(m-1)\,.\, s_b'^2 - \sum_1^m \frac{N_i^2}{N_0^2}\,(1-f_i) \cdot \frac{s_i^2}{n_i} \right] + \frac{v\ (\hat{Y})}{M\,(M-1)\,N_0^2}$$

என்று பெறப்படுகிறது.

இதில் $s_b'^2 = \frac{1}{m-1} \sum_1^m \left(\frac{N_i}{N_0} \cdot \bar{y}_i - \bar{Y} \right)^2$ ஆகும்.

$$\left[\hat{\bar{Y}} = \frac{\hat{Y}}{MN_0} \right].$$

எனவே, $V\ (\hat{Y})$-ன் மதிப்பீடு,

$$v\ (\hat{Y}) = M^2\,.\,N_0^2\,.\,(1-f)\,.\,\frac{s_b'^2}{m} + \frac{M}{m} \sum_1^m N_i^2\,.\,(1-f_i)\,\frac{s_i^2}{n_i}$$

என்றபடி அமைகிறது.

{ குறிப்பு : இயல்பாக அமையப்பெற்ற பல்வேறு பிரிவுகளைக் கருதுகையில், முதல் கட்டத்தில் தேர்ந்தெடுக்கும் திரள்களின் அளவுகள் (cluster sizes) மாறுபடும் தன்மையுடன் இருக்கும். உதாரணமாக, குடும்பங்களைக் கருதுகையில், நபர்களின் எண்ணிக்கையானது ஒவ்வொரு குடும்பத்துக்கும் சமமானதாக அமைவதில்லை. திரள்களாகக் கருதப்படும் கிராமங்களும் இத்தன்மையதே.}

பின்குறிப்பு :

திரள் மாதிரி முறை ஆய்வு முழுவதிலும், நாம் கருதிய திரள்கள் அனைத்துமே சம அளவினதாக (equal size)

அமைந்தன. எனவே, மேற்கண்ட மதிப்பீடுகள் [$\hat{Y}$, $V(\hat{Y})$ போன்றவை] ஒரு பொதுவான நிலையினைப்பற்றியதாய் அமைகின்றன. இதில், ஒரு குறிப்பிட்ட நிலைதான் சம அளவினதான (equal size) திரள்களைக் கருதும் முறையாகும்.

இனி, இரு நிலைகளிலும் (first and second stages) திரும்ப வைக்கும் சம வாய்ப்பு முறையினைக் (SRSWR) கருதி முடிவுகளைக் காணலாம் :

இரு கட்டங்களிலும், திரும்ப வைக்கும் முறை பின்பற்றப் படுகையில், $\hat{Y}$ மாறுவதில்லை.

$$V(\hat{Y}) = M^2.N_0^2 \left[\frac{\sigma_b^2}{m} + \frac{\sigma_w^2}{mn}\right]$$

என்று அமையும்.

$$\sigma_w^2 = \frac{1}{M}\sum_{M} \frac{N_i}{N_0}.\ \sigma_{wi}^2 \ ; \ \sigma_{wi}^2 = \frac{N_i - 1}{N_i}.\ S_i^2$$

(உள்ளீட்டு மாறுபாடு) $\left[f_i = \frac{n}{N_0}\right.$.

மேலும்,

$$\sigma_b^2 = \frac{\sigma^2}{N}\ \{1 + \overline{M-1}.\ \rho_c\}$$ என்றும் காணலாம்.

இதில்,

$$\rho_c = \frac{\sum\limits_{i}\sum\limits_{j}\sum\limits_{j \pm j'} (Y_{ij} - \bar{Y})\ (Y_{ij'} - \bar{Y})}{MN\ (N-1)\ \sigma^2}$$ என்றாகும்.

பிரிவு உள்ளீட்டு ஒட்டுறவுக்கெழு (intraclass correlation). முழுமைத் தொகுதி சராசரியின் மதிப்பீடு :

$$\bar{Y} = \frac{Y}{M.N_0} \quad \left(N_0 = \frac{1}{M}\ \Sigma N_i \right)$$

என்ற சராசரியின் மாதிரி மதிப்பீடு, $\hat{Y}$-ஐ $M.N_0$-ஆல் வகுத்துப் பெறப்படும். இதற்கு N_0 மதிப்பு தெரிந்திருக்கவேண்டும். அந் நிலையில், மாறுபாட்டு மதிப்பீடும் பெறப்படும். ஆனால் N_0 மதிப்பு தெரியாத நிலையில், அதன் மதிப்பினை, m அளவு முதற் கூறுகளினின்று வரும் துணைக் கூறுகளின் எண்ணிக்கைச் சராசரி மூலம் மதிப்பிட்டு அறிதல் வேண்டும். அதற்கு,

$$\hat{Y} = \frac{M}{m} \sum_{1}^{m} \frac{N_i}{n_i} \sum_{n_i} y_{ij} \Bigg/ \frac{M}{m} \sum_{1}^{m} N_i$$

என்று அமையும். (இது பிறழ்ச்சியுடைய மதிப்பீடாகும்)

மூன்றுகட்ட மாதிரி முறை
(Three stage Sampling Method)

இரு கட்ட மாதிரி முறையானது மேலும் விரிவாக்கப்படுகிறது. முதல் கட்டத்தில் கல்லூரிகளைத் தெரிவு செய்து, அடுத்த கட்டத்தில் வகுப்புகளைக் கருதி அந்த அடிப்படையில் மாதிரியினை எடுத்து, மூன்றாவது நிலையில் (stage) இறுதிக் கூறுகளான மாணவர்களைக் கருதுவது இதற்கு ஓர் எடுத்துக் காட்டாகும். மாணவர் தம் பிரச்சினை ஏதேனும் பற்றி ஆய்வு நடத்தப்பட வேண்டின், இம் முறை கையாளத்தக்கது. விவசாயத்துறை ஆய்வுகளில், பயிர் விளைவு பற்றி ஆய்வு நடத்தப்பட, முதலில் கிராமங்களைத் தேர்ந்தெடுத்து, அதன்பின் நிலங்களைத் தெரிவு செய்து, மூன்றாவதாகப் பயிரிடும் சிறு நிலப்பரப்புகள் கருதப்படுகின்றன. (*NSS* ஆய்வுகளில் பெரும்பாலும் இந் நிலையானது அமைந்திடும்).

முதல் கட்டத்தில் L திரள்களை (முதற் கூறுகள்) எடுத்துக் கொண்டு, அடுத்து M 2-ம் நிலைக் கூறுகளைக் கருதி, இறுதிநிலையில் N கூறுகள் (முடிவுநிலை) கருதப்பட்டு, ஒவ்வொரு கட்டமாக மாதிரிகள் எடுக்கப்படுகின்றன. l, m, n என்பவை முறையே ஒவ்வொரு கட்டத்திலும் எடுக்கப்படும் மாதிரி அளவுகளாகும்

y_{ijk} என்பது இறுதியில் பெறப்படும் மதிப்பு.

Y_{ijk}-யினை முழுமைதொகுதிக் கூறாகக் கருதலாம்.

செயல்முறை எளிதாக அமையவேண்டுமெனில், இம் முறையில், மாறுபடும் வாய்ப்பு மாதிரி முறையைப் பின்பற்றுவது சிறந்தது எனலாம். அதனைக் கருதும் முன்பாக, சில மதிப்பீடுகளை, ஏற்கெனவே ஆராய்ந்ததன் அடிப்படையில் பெறுகிறோம்.

முழுமைத் தொகுதி மொத்தம் மற்றும் சராசரி இவற்றினை மதிப்பிடல் :

முதல் கட்டத்தில் தெரிந்தெடுக்கப்படும் l கூறுகள் ஒவ்வொன்றிலும் m' அளவு மாதிரி எடுக்கப்பட்டு, மூன்றும் கட்டத்தில் n_{ij} என்றபடி இறுதிநிலைக் கூறுகள் தேர்ந்தெடுக்கப்படுவதாகக் கொள்க.

அப்பொழுது,

$$\hat{Y} = \frac{L}{l} \sum_{L} \frac{M_i}{m'} \sum_{j=1}^{m'} \frac{N_{ij}}{n_{ij}} \sum_{k=1}^{n_{ij}} y_{ijk}$$

என்பது Y-ன் மாதிரி மதிப்பீடாகும்.

(இது ஒரு பிறழ்ச்சியற்ற மதிப்பீடு.)

L முதல் நிலைக் கூறுகளின் சராசரி $\bar{Y}$ எனக்கொண்டு,

$\hat{Y} = N . \hat{\bar{Y}}$ என்ற அடிப்படையைப் பயன்படுத்த,

$\hat{Y} = L . \hat{\bar{Y}}$ என்று பெறலாம்.

இது

$\hat{Y} = L \left[\frac{1}{l} \Sigma \hat{Y}_i \right]$ என்று ஆகும்.

அதாவது

$$\hat{Y} = \frac{L}{l} \left[\sum_{i} M_i . \hat{\bar{Y}}_i \right] \text{ (அடுத்த கட்டம்).}$$

$$= \frac{L}{l} \left[\sum_{i} \frac{M_i}{m'} \sum_{j} \hat{Y}_{ij} \right]$$

$$= \frac{L}{l}\left[\sum_{i}\frac{M_i}{m'}\sum_{j}N_{ij}\ \bar{y}_{ij}\right] \text{(மூன்றாவது நிலை)}$$

$$\therefore\ \hat{Y} = \frac{L}{l}\left[\Sigma\ \frac{M_i}{m'}\sum_{j}\frac{N_{ij}}{n_{ij}}\sum_{1}^{n_{ij}} y_{ijk}\right]$$

என்று அமைகிறது. இதிலிருந்து,

$$\hat{\bar{Y}} = \frac{\hat{Y}}{N} \text{ என்று பெறலாம்.}$$

$\hat{Y}$-ன் மாறுபாட்டு அளவு (Variance of $\hat{Y}$) :

இரண்டு இருகட்ட திரள்மாதிரித் தேர்வு மாதிரிமுறைசார்ந்த மாறுபாடுகளை (sampling variances) கருதி, மூன்று கட்ட மாதிரித் தேர்வில், $V(\hat{Y})$ ஆனது மதிப்பிடப்படும்.

இரு கட்ட நிலையில்

$$V(\hat{Y}) = M^2.\frac{(M-m)}{M}.\ \frac{S_b^2}{m}$$

$$+\ \frac{M}{m}\sum_{M} N_i^2.\ \frac{(N_i - n_i)}{N_i}.\ \frac{S_i^2}{n_i}$$

என்று பெற்றோம்.

இதில்,

$$S_b^2 = \frac{1}{M-1}\sum_{1}^{M}(Y_i - \bar{Y})^2 \qquad \text{(இடைப்பட்ட மாறுபாடு)}$$

$$S_i^2 = \frac{1}{N_i - 1}\sum_{1}^{M}(Y_{ij} - \bar{Y}_i)^2 \qquad \text{(உள்ளீட்டு மாறுபாடு).}$$

இதனை, முதல் இரு கட்டங்களுக்கு, மூன்று கட்டத் தேர்வு அளவில் கருத,

$$L^2 \cdot \frac{(L-l)}{L} \cdot \frac{S_b^2}{l} + \frac{L}{l} \cdot \sum_{1}^{L} M_i^2 \cdot \frac{M_i - m'}{M_i} \cdot \frac{S_i^2}{m'}$$

என்று பெறுவோம்.

இதேபோல், இரண்டாவது, மற்றும் மூன்ருவது நிலைகளைக் (கட்டங்களை) கருதி,

$$M_i^2 \cdot \frac{M_i - m'}{M_i} \cdot \frac{S_i^2}{m'} + \frac{M_i}{m'} \sum_{1}^{M_i} N_{ij}^2 \cdot \frac{N_{ij} - n_{ij}}{N_{ij}} \cdot \frac{S_{ij}^2}{n_{ij}}.$$

என்று பெறுவோம்.

$$\left[S_{ij}^2 = \frac{1}{N_{ij}-1} \sum_{k=1}^{N_{ij}} \left(Y_{ijk} - \overline{\overline{Y}}_{ij} \right)^2 \text{ ஆகும் } \right].$$

மூன்று கட்ட மாதிரித் தேர்வில்,

$$V(\hat{Y}) = L^2 \cdot \frac{L-l}{L} \cdot \frac{S_b^2}{l} + \frac{L}{l} \sum_{L} \left[\left(M_i^2 \cdot \frac{M_i - m'}{M_i} \cdot \frac{S_i^2}{m'} \right) + \frac{M_i}{m'} \cdot \sum_{M_i} N_{ij} \cdot \frac{N_{ij} - n_{ij}}{N_{ij}} \cdot \frac{S_{ij}^2}{n_{ij}} \right]$$

$$\therefore \quad V(\hat{Y}) = L^2 \cdot \frac{L-l}{L} \cdot \frac{S_b^2}{l} + \frac{L}{l} \sum_{L} M_i^2 \cdot \frac{M_i - m'}{M_i} \cdot \frac{S_i^2}{m'} + \frac{L}{l} \sum_{L} \frac{M_i}{m'} \sum_{M_i} N_{ij}^2 \cdot \frac{N_{ij} - n_{ij}}{N_{ij}} \cdot \frac{S_{ij}^2}{n_{ij}}$$

என்று பெறப்படும்.

{ மூன்று நிலைகளிலும் கூறுகளின் தேர்வு ஒரே மாதிரியான செலவில் இருப்பின், அதிக அளவில் முதல்நிலைக் கூறுகளை எடுப்பதன்மூலம் $V(\hat{Y})$-ன் அளவினைக் குறைக்க இயலும். } மற்ருெரு முடிவு :

$V(\hat{Y})$-ன் மாதிரி மதிப்பீடு :

$$\hat{V}(\hat{Y}) = M^2 \cdot \frac{M-m}{M} \cdot \frac{S_b{}^2}{m} + \frac{M}{m} \sum_m N_i{}^2 \cdot \frac{N_i - n_i}{N_i} \cdot \frac{s_i{}^2}{n_i}$$

இதில்,

$$s_b{}^2 = \frac{1}{m-1} \sum_m (\hat{Y}_i - \hat{\bar{Y}})^2 ;$$

$$s_i{}^2 = \frac{1}{n_i - 1} \sum_{n_i} (y_{ij} - \bar{y}_i)^2 ;$$

$$\hat{Y}_i = \frac{N_i}{n_i} \sum_{n_i} y_{ij} ; \quad \hat{\bar{Y}} = \frac{1}{m} \sum_m \hat{Y}_i ;$$

$$\bar{y}_i = \frac{y_i}{n_i} ; \; y_i = \sum_{n_i} y_{ij}$$

என்று அமைந்திருக்கும்.

* *PPS* முறை (ஒவ்வொரு கட்டத்திலும்) :

திரும்பவைக்கும் முறையைப் பின்பற்றி, அளவோடு விகிதத் தொடர்பில் வாய்ப்புகளை (மாறுபடும் வாய்ப்பு) கருதி, மாதிரிகள் எடுக்கப்படுகின்றன. $\{P_i\}$; $\{P_{ij}\}$; $\{P_{ijk}\}$ என்பவை ஒவ்வொரு கட்டத்திலும் அமையும் வாய்ப்பு நிலைகள். இந்நிலையில்,

$$\hat{Y} = \frac{1}{lmn} \sum_{1}^{l} \frac{1}{p_i} \sum_{j=1}^{m} \frac{1}{p_{ij}} \sum_{k=1}^{n} \frac{y_{ijk}}{p_{ijk}}$$

என்று பெறப்படும். (பிறழ்ச்சியற்ற மதிப்பீடு).

மேலும்,

$$E(\hat{Y}) = E_1 E_2 E_3 (\hat{Y}) \quad \text{(மூன்று நிலைகளைக் கருதி).}$$

* அடுத்த அத்தியாயத்தில் விரிவாக விளக்கப்படுகிறது.

$$V(\hat{Y}) = V_1 \,.\, E_2E_3(\hat{Y}) + E_1 \quad V_2 \,.\, E_3 \quad \hat{Y})$$
$$+ E_1E_2 \,.\, V_3(\hat{Y}))$$

$\hat{Y}$-ன் மாறுபாட்டு அளவு :

$$V(\hat{Y}) = \frac{1}{l}\left(\sum_{1}^{L} \frac{Y_i^2}{P_i} - Y^2\right)$$
$$+ \frac{1}{lm}\sum_{1}^{L}\frac{1}{P_i}\left(\sum_{1}^{M_i}\frac{Y_{ij}^2}{P_{ij}} - Y_i^2\right)$$
$$+ \frac{1}{lmn}\sum_{1}^{L}\frac{1}{P_i}\sum_{1}^{M_i}\frac{1}{P_{ij}}\cdot\left[\sum_{1}^{N_{ij}}\frac{Y_{ij}^2}{P_{ij}} - Y_{ij}^2\right]$$

முதல் கட்டத்தில், திரும்பவைக்கும் முறை பயன்படுத்தப் படுவதால், n-பிறழ்ச்சியற்ற மதிப்பீடுகளான

$$\frac{1}{p_i}\cdot\frac{1}{m}\sum_{1}^{m}\frac{1}{p_{ij}}\left[\frac{1}{n}\sum_{1}^{n}\frac{y_{ijk}}{p_{ijk}}\right] \;\ldots\ldots\; (A)$$

அனைத்துமே தொடர்பற்றும், ஒரே மாறுபாட்டு அளவுடனும் அமைகின்றன.

[$\therefore \; v(\hat{Y}) = \frac{1}{l(l-1)}\sum_{1}^{l} t_i^2 - n\bar{t}^2$ (t_i என்பது

(A) மதிப்பாகும்.)]

இங்கு, மாறுபாடானது

$V(\hat{Y}) = \frac{A_1}{l} + \frac{A_2}{lm} + \frac{A_3}{lmn}$ என்ற முறைப்படி அமையும்.

(A_1, A_2, A_3 முன்பே கொடுக்கப்பட்டவை.)

IV. பல்தோற்ற மாதிரிமுறை :
(Multi-Phase Sampling Method)

பலகட்டத் தேர்வு முறையினின்று மாறுபடும் இத் தேர்வு முறை 'பல்தோற்ற' மாதிரி எனவும் கூறப்படும். இதன்படி, முதல் தோற்றமாக (first phase), ஏதேனும் துணை இயல்புகளை (auxiliary characteristic) பற்றிய அளவில் முதலில் கூறுகள் தெரிவு செய்யப்பட்டு, பின்பு ஏற்கெனவே எடுக்கப்பட்ட மாதிரியி னின்று உட் கூறுகளை எடுத்து, முதலில் கருதப்பட்ட துணை இயல்பினைப் பயன்படுத்தி, படுகைகள் அமைத்து வரிசைப்படுத்தி, தேவைப்படும் நோக்கத்திற்கேற்ப மதிப்பீடுகளைப் பெறலாம். இதனைப் பல தோற்றமாகக் கருதுகையில், பல்தோற்ற மாதிரி என்கிறோம்.

பலகட்டத் தேர்வில், ஒவ்வொரு கட்டத்திலும் (stage), அடுத்த கட்டத்திற்கு எடுக்கப்பட வேண்டிய கூறுகள் அமைந்த திரள்களை (clnsters) தெரிந்தெடுக்கிறோம். இதற்கு மாறாக, பல்தோற்ற மாதிரிமுறையில் முதல் நிலையினைத்தவிர, ஒவ்வொரு நிலையிலும் முன்பு எடுக்கப்பட்ட அதே கூறுகளின் உட்கூறுகள் துணைமாதிரி வடிவில் (subsamples) எடுக்கப் படுகின்றன. எனவே, இத்தகைய வழிமுறையில், இறுதிநிலைக் கூறுகளைப் பற்றிய முழுவிவரமும் கொடுக்கப்படல் வேண்டும். [ஆனால், பலகட்டத் தேர்வில் இத்தகைய மாதிரி சட்ட (sampling frame) வரையறை, ஏற்கெனவே தெரிவு செய்யப் பட்ட திரள்களினைப் பொறுத்த அளவில் அமைவதே போதுமான தாகும்.] இதன்மூலம் ஆய்வுச் செலவு பெரும் அளவில் குறைக்கப்படுகிறது.

ஒரு பிரிவு மக்களின் செலவுபற்றிய ஆய்வில், ஏற்கெனவே பெறப்படும் விவரப்பட்டியலை வைத்துக்கொண்டு, இதோடு தொடர்புடைய வேறு பல இயல்புகளைப் பற்றியும் (auxiliary) விவரங்களைப் பெற்று (குடும்ப நபர்கள் எண்ணிக்கை, பெறும் வருமானம், வாழ்வு வசதிகள் போன்ற விவரங்கள்), அவற்றை அடுத்த நிலை ஆய்வினில் பயன்படுத்தலாம். இதில் ஒவ்வொரு தோற்றத்திலும் (phase) பெறப்படும் மதிப்பீடுகள், அடுத்த முகத்தில் எடுக்கப்படும் துணைமாதிரி (subsample) கூறின் அடிப்படையில் அமைந்திருக்கும் என்பதுகுறிப்பிடத் தக்கது. எனவே இம்முறை, பலகட்ட மாதிரிமுறையினின்று மாறுபடுவதை உணரலாம்.

பலகட்ட மாதிரித் தேர்வு முறையில், ஒவ்வொரு கட்டத்திலும் வெவ்வேறு வகைப்பட்ட மாதிரிக் கூறுகளை (units) தேர்ந்தெடுப்பதைக் கருதுகிறோம். இதற்கு மாறாக பல தோற்ற அல்லது பன்முக மாதிரித் தேர்வு முறையினில், ஒரே வகையான மாதிரிக்கூறுகள் ஆய்வின் பல தோற்றங்களிலும் தேர்ந்தெடுக்கப்படுகின்றன. குறிப்பாக, முதல் தோற்றத்தில் ஒரு முழு அளவில் மாதிரியானது எடுக்கப்படுகிறது. அதன்பின் அதன் கூறுகளின் சில இயல்புகள் ஆராயப்படும். பிறகு இரண்டாவது தோற்றத்தில் அம் முழு மாதிரியினின்று ஏதேனும் துணை மாதிரியை எடுத்து அதன் கூறுகளின் வேறு சில இயல்புகள் ஆராயப்படும். உதாரணமாக, தேர்ந்தெடுக்கப்பட்ட 1000 குடும்பங்களின் வருமானம் பற்றிய விவரங்களைத் தொகுப்பதாகக் கொள்வோம். அக் குடும்பங்களில், ஏதேனும் வழியில் 100 குடும்பங்களைத் துணைமாதிரியாகத் தேர்ந்தெடுத்து, அவற்றின் வருமான விவரங்களை மட்டுமன்றி, வருமானம் ஈட்டப்பட்ட வழிகள், அவ் வருமானம் எவ்வாறு செலவிடப்படுகிறது, குடும்ப நபர்கள் விவரம் போன்ற இதர விவரங்களையும் திரட்டும் நிலையில் இரு தோற்ற மாதிரிமுறை அமைகிறது. இது 'இரு மாதிரி' முறை எனவும் கூறப்படும் (Double Sampling). முதல் தோற்றத்தில் (முகத்தில்) விவரங்கள் சேகரிக்கப்படுகையிலேயே, இரண்டாம் தோற்றத்தில் இடம் பெறும் விவரங்களைத் திரட்டலாம். அல்லது தனியாகக் கருதியும் திரட்டலாம். இத்தகைய தோற்றங்கள் வேறு வேறு காலங்களில் திரட்டப்படுவனவாகவும் அமையக்கூடும்.

முதல் தோற்றத்திலேயே ஒரு விரிவான மாதிரியினைத் தெரிந்தெடுத்து, அதன் கூறுகளைக் கருதித் தேவைப்பட்ட விவரங்களைப் பெறலாம். அதிக விவரங்கள் இல்லாதபடியினால் இதற்கு ஆகும் செலவும் குறைவாகும். பின்பு, இதன்மூலம் பெறப்படும் முடிவுகளை வைத்து, முதலில் எடுக்கப்படும் மாதிரியினின்று சிறு அளவில் துணை மாதிரியினை தேர்ந்தெடுத்து, அதனின்று முதல் தோற்றத்தில் கருதப்படும் விவரங்களையே அதிக கவனத்துடன் கூர்ந்து ஆராயலாம். இரண்டாம் தோற்ற மாதிரியானது, முதலில் எடுக்கப்பட்ட மாதிரியின் உட்பிரிவாகவே அமையவும் நேரிடலாம் (ஆனால் இத்தகைய நிலை கட்டாயமானதல்ல). சில நேரங்களில், முக்கிய நோக்கமுள்ள மாறியினைப் பற்றிய விவரங்களை இரண்டாம் தோற்றத்தில்மட்டுமே திரட்டும் நிலையும் உண்டு. இதற்கு, முதல் தோற்ற நிலையில், அம் மாறியோடு தொடர்புடைய ஒரு துணை மாறியினை அமைத்துக் கொள்வதன் மூலம், அதன் விவரங்களை சேகரித்து, அந்த ஆராய்ச்சியின்

அடிப்படையில், இரண்டாவது தோற்றத்துக்கு வேண்டிய படுகைகளை அல்லது பிரிவுகளை அமைத்து, அதன்பின் மாதிரி களைத் தேர்ந்தெடுத்து, குறிப்பிட்ட மாறியின் இயல்புகளைப் பற்றிய விரிவான ஆய்வு நடத்தப்படலாம்.

மதிப்பீடுகள் :

இருதோற்ற மாதிரி முறையினைக் கருதுவோம். இது 'இருமாதிரி' முறையாகக் கருதப்படும் (Double Sampling).

முதலில், m_1 கூறுகள் கொண்ட மாதிரியினை ஈடுசெய்யப் படாத சமவாய்ப்பு முறைப்படி தெரிவு செய்து, அடுத்த நிலையில் அதனின்று m_2 கூறுகள் கொண்ட $(m_2 < m_1)$ துணை மாதிரியினை (subsample) திரும்ப வைக்கும் முறையில், PPS முறை மூலமாக (அளவோடு விகிதத் தொடர்புடைய வாய்ப்புகளுடன்) தெரிவு செய்யப்படுவதாகக் கொள்வோம்.

முழுமைத் தொகுதி மொத்தமான Y-ன் பிறழ்ச்சியற்ற மதிப்பீடு

$$\hat{Y} = \frac{M}{m_1} \cdot \frac{1}{m_2} \sum_{1}^{n_2} \frac{y_j}{p_j}$$ ஆகும்.

(M : முழுமைத் தொகுதியின் அளவு).

இதில் $p_j = \frac{x_j}{x'}$ என்று காணப்படும்..

y_j என்பது, துணைமாதிரியின் j-ஆவது உறுப்பின் y மதிப்பு ; x' என்பது முதல் தோற்றத்தில் எடுக்கப்படும் மாதிரியின் x-ன் மொத்தம் (x என்பது துணை இயல்பினைக் குறிக்கும் மாறியாகும்.

மேலும்,

$$E(\hat{Y}) = E_1 \, E_2 = (\hat{Y})$$

$$= E_1 \left\{ \frac{M}{m_1} \sum_{m_1} y_i \right\} = Y$$ என்று காணலாம்.

மாறுபாட்டு அளவு

$$V(\hat{Y}) = V_1 \, E_2 \, (\hat{Y}) + E_1 \, V_2 \, (\hat{Y})$$

என்று வரையறுக்கப்பட்டு,

$$V(\hat{Y}) = V_1 \left\{ \frac{M}{m_1} \sum_{m_1} y_i \right\}$$

$$+ E_1 \left\{ \frac{M^2}{m_1^2} \cdot \frac{1}{m_2} \right\} \left\{ x' \sum_{m_1} \frac{y_i^2}{x_i} - \left(\sum_{m_1} y_i \right)^2 \right\}$$

$$\therefore \quad V(\hat{Y}) = M^2 \cdot \frac{(M - m_1)}{M - 1} \cdot \frac{\sigma_y^2}{n_1}$$

$$+ \frac{M}{M-1} \cdot \frac{m_1 - 1}{m_1} \cdot \frac{V_{pps}}{m_2}$$

என்பதாகப் பெறப்படும்.

V_{pps} என்பது, மொத்த முழுமைத் தொகுதியினின்று, PPS முறையில் தெரிவு செய்யப்படும் ஒரு மாதிரிக் கூறின் அடிப்படையில் பெறப்படும், Y மதிப்பீட்டின் மாறுபாட்டு அளவாகும்.

$\frac{m_1 - 1}{m_1}$ என்பதை '1' என்று கொண்டு (m_1 பெரிதாக இருக்கையில்,

$$V(\hat{Y}) = \frac{A_1}{m_1} + \frac{A_2}{m_2} + A_3$$

என்ற வடிவில் எழுதப்படும்.

இந்த வரையறையில் $\frac{1}{m_1}, \frac{1}{m_2}$ இவற்றின் கெழுக்களான A_1, A_2, மற்றும் மாறிலி A_3 இவை பின்வருமாறு :—

$$A_1 = \frac{1}{(M-1)} \cdot \sum_{1}^{M} \left(\frac{N_i \bar{Y}_i}{N_0} - \bar{Y} \right)^2$$

$$- \frac{1}{M} \sum_{1}^{M} \frac{1}{N_i} \left(\frac{N_i}{N_0} \right)^2 ;$$

$$A_2 = \frac{1}{M} \sum_{1}^{M} \left(\frac{N}{N_0}\right)^2 \Big/ \left(\frac{n_i}{n}\right); \qquad \left[N_0 = \frac{1}{M} \sum_{1}^{M} N_i\right]$$

$$A_3 = -\frac{1}{M} \cdot \frac{1}{(M-1)} \cdot \sum_{1}^{M} \left(\frac{N_i \bar{Y}_i}{N_0} - \bar{Y}\right)^2$$

என்று மதிப்பிடப்படுகின்றன.

பன்முகத் தேர்வு முறையில், செலவுச் சார்பு

$$c = c_0 + c_1 \, . \, m_1 + c_2 \, . \, m_2$$

என்று கொள்ளப்படுகிறது. c_1 என்பது துணை மாறியின் (auxiliary variable) மதிப்பினைக் கண்டறிந்து பயன்படுத்த ஆகும் செலவு வீதமாகவும், c_2 என்பது கருதப்படும் முக்கிய மாறியின் மதிப்பைக் கண்டறிந்து பயன்படுத்த ஆகும் செலவு வீதமாகவும் அமையும். $c_2 > c_1$ என்பதாகவும் காணப்படும்.

m_1, m_2 இவற்றின் உத்தம மதிப்புகள் கொடுக்கப்பட்ட செலவில், மாறுபாட்டினைச் சிறும அளவினதாக்கத் தக்கவிதத்தில் கணக்கிடப்படும். அவை

$$m_1 = \frac{(c' - c_0) \, . \, \sqrt{A_1/c_1}}{\sqrt{A_1 \, . \, c_1} + \sqrt{A_2 \, . \, c_2}}$$

என்றும், $$m_2 = \frac{(c' - c_0) \, . \, \sqrt{A_2/c_2}}{\sqrt{A_1 \cdot c_1} + \sqrt{A_2 \, . \, c_2}}$$

என்றும் அமையும்.

கொடுக்கப்படும் மாறுபாட்டு அளவிற்கேற்ப (V_0), செலவினைச் சிறும அளவினதாக்கும் மதிப்புகள்

$$m_1 = \frac{\sqrt{A_1 c_1} + \sqrt{A_2 c_2}}{(V_0 - A_3)} \cdot \sqrt{\frac{A_1}{c_1}}$$

$$m_2 = \frac{\sqrt{A_1 c_1} + \sqrt{A_2 c_2}}{V_0 - A_3} \cdot \sqrt{\frac{A_2}{c_2}}$$

என்று கணக்கிடப்படுகின்றன.

8. மாறுபடும் வாய்ப்பு மாதிரித்தேர்வு (அல்லது) PPS மாதிரிமுறை

(Varying Probability Sampling)

அறிமுகம்—சராசரி மற்றும் மொத்த மதிப்பீடு—$\bar{y}_{pps}$-ன் மாறுபாடு—PPS முறை, எளிய திரள் முறை—ஓர் ஒப்பீடு—எடுத்துக்காட்டுகள், பயிற்சிகள்—PPS முறையில் மாதிரிப் பங்கீடு.

அறிமுகம்

முன்பு கருதிய திரள் மாதிரித் தேர்வுமுறையில் (cluster sampling) மொத்தம் உள்ள M திரள்களிலிருந்து முதல் கட்ட மாதிரித் தேர்வு நடைபெறும்பொழுது, ஒவ்வொரு திரளும் தெரிந்தெடுக்கப்பட, $\frac{1}{M}$ என்ற சம வாய்ப்பினை (equal probability) பெறுவதாகக் கருதப்பட்டது. உதாரணமாக, ஓர் ஆய்வில் 10 கிராமங்கள் கருதப்படின், ஒவ்வொரு கிராமமும் மாதிரியில் இடம் பெற $\frac{1}{10}$ என்றபடி சமவாய்ப்புப் பெறுகிறது. இம் மாதிரியான சமவாய்ப்பு முறை, சில நிலைகளில் சரியான வழிமுறையாக இருப்பதில்லை. எனவே, திரள் மாதிரித் தேர்வு முறையானது, மாறுபடும் வாய்ப்புடன் (varying probability) கூடிய ஒரு நிலையினை இங்கு விரிவாகக் கருதுவோம்.

எடுத்துக்காட்டாக, ஒவ்வொரு குடும்பத்துக்கும் துணி வகைக்குச் சராசரியாக ஆகும் செலவு பற்றிய ஆய்வில், முதல் கட்டத் தெரிவிற்கு, 100 நகர்களைத் திரள்களெனக் கருதி, அவற்றுள் 10 நகர்கள் தேர்ந்தெடுக்கப்படுவதாகக் கொள்வோம். இரண்டாம் இறுதிக்கட்டத்தில், இத் திரள்களினின்று n_i-என்ற அளவில் குடும்பங்களைத் தேர்ந்தெடுக்கும் படியாக அமைந்திட்டால், பெரிய அளவினதான நகர்களிலிருந்து அதிக வீதத்திலும், சிறிய நகர்களிலிருந்து குறைந்த வீதத்திலும் மாதிரிகள் எடுக்கப்படின், குடும்பங்களைத் தெரித்தெடுக்க ஏற்படும் வாய்ப்புகள் சமநிலையில் அமையக் காணலாம்.

அதாவது, தேர்ந்தெடுக்கப்படும் மாதிரியானது, முழுமைத் தொகுதியின் தக்க பிரதிநிதியாக அமையவேண்டுமெனில், திரள்கள் ஒவ்வொன்றின் அளவும் (size) வெவ்வேறாக இருக்கும் நிலையில்,பயன்படுத்தப்படும் வழிமுறையானது அத் தொகுதியின் ஒவ்வோர் உறுப்பும் தெரிந்தெடுக்கப்பட சம வாய்ப்பினை அளிப்பதாக இருக்கவேண்டும். இதற்கேற்றபடி, சில திரள்கள் பெரிய அளவினதாக இருப்பின், அவை தெரிந்தெடுக்க அதிக வாய்ப்பினைப் பெறும்படியாகவும், குறைந்த அளவினதாக இருப்பின், வாய்ப்புக் குறைந்தும் அமைதல் வேண்டும். இதன் மூலம், பல்வேறு திரள்களினின்று தெரிந்தெடுக்கப்படும் மாதிரிக் கூறுகளின் ' தெரிவு பெறும் ' வாய்ப்பானது சம அளவில் (equal) அமைய ஏதுவாகிறது.

எனவே, முதல்நிலைக் கூறுகளான திரள்களின் அளவிற்கேற்ற விதத்தில் (proportional to size) வாய்ப்புகளை நிர்ணயித்து மாதிரி எடுக்கும் முறையானது, ' திரள் அளவு விகிதத்தில் வாய்ப்புடைய' மாதிரிமுறை (sampling with probability proportional to size) எனப்படுகிறது. சுருக்கமாக இதனை ' PPS ' முறை என்று குறிப்பிடலாம். பிறழ்ச்சியற்ற மதிப்பீடுகளைத் தரவும், மற்ற முறைகளைவிட அதிக அளவில் திட்டமுடைய மதிப்பீடுகளைப் பெறவும் வல்லதொரு முறையாக இது அமைவதைக் காணலாம்.

திரள் அளவின் (cluster size) வீதத்தில் வாய்ப்புப் பெறும் மாதிரியானது, பின்வரும் எடுத்துக்காட்டில் விளக்கப்படுகிறது :

திரள் எண்	அளவு (N_i)	ΣN_i
1	6	6
2	5	11
3	8	19
4	15	34
5	16	50

முதல்திரளினைத் தெரிவு செய்யும் வாய்ப்பு $\frac{6}{50}$, 2-வது திரளிற்கு $\frac{11}{50}$ என்றபடியாக வாய்ப்புநிலைகள் அமையக் காணலாம். இவற்றைத் தெரிவு செய்து மாதிரியினை நிர்ணயிப்பதானது,

ராண்டம் எண்களைப் பயன்படுத்திப் பெறப்படும். 1 முதல் 6 வரையில் உள்ள எண்களில் ஏதேனும் ஒன்று தெரிவு செய்யப் பெறின், முதல் திரள் ; 7-லிருந்து 11 வரை 2 வது திரள், என்ற வகையில் மாதிரிக் கூறுகள் தேர்ந்தெடுக்கப்படும். திரும்ப வைக்கும் (with replacement) முறையினைப் பயன்படுத்துவது, மதிப்பீடுகளைப் பெற எளிதான வழியாகும் இது முதல் கட்டத் தேர்வாகும்.

இரண்டாம் கட்டத்தில், துணைமாதிரிக் கூறுகளை (sub samples), (i) திரும்ப வைக்கும் முறை அல்லது (ii) திரும்ப வைக்காத முறையின்மூலம் தெரிந்தெடுக்கலாம். ஒரே திரள் கூறானது (primary unit) திரும்பவும் எடுக்கப்பட்டு, புதியதொரு துணைக்கூறு (subsample) தெரிவு பெறின், திரும்பவைக்காத முறைப்படி அமையும் நிலையிலும், ஏற்கெனவே தெரிந்தெடுக்கப் பட்ட முதல்கூறின் உட்கூறுகச் சில உறுப்புகள் அமைய நேரிடும். எனவே, இரண்டாவது துணைமாதிரியின் சில உறுப்புகள் முதலாவது துணைமாதிரியின் உறுப்புகளாக அமையக் காணலாம். இத் துணைமாதிரிகள் ஒன்றேடொன்று தொடர் பற்றவை.

இங்கு, திரும்பவைக்காத முறையின்மூலம் இரண்டாம் கட்டத் தேர்வு நடத்தப்படும் நிலையினைக் கருதுகிறோம்.

∴ துணைமாதிரிக் கூறுகள் ஒன்றேடொன்று தொடர்பற்று அமையும்.

முழுமைத் தொகுதியின் சராசரி மற்றும் மொத்த அளவு (total) இவற்றின் மதிப்பீடு

சுய-எடையிடப்பட்ட(self-weighted) மாதிரிகளில், ஏதேனும் ஒரு துணைக்கூறானது(secondary unit)மாதிரியில் தெரிந்தெடுக்கப் படும் வாய்ப்பானது எல்லாத் தொகுதிக் கூறுகளுக்கும் சமமான அளவினதாக இருக்கும். (தனியான எந்தக் காரணியினையும் பயன்படுத்தி எடையிடப்படாத மாதிரி, சுய எடையிட்ட மாதிரியாகும்).

எனவே, ஒவ்வொரு துணைக்கூறும் n-அளவுள்ள மாதிரியில் இடம்பெற $\frac{n}{N}$ என்னும் சமவாய்ப்பினைப் பெறும். m-முதற் கூறுகளை M கூறுகள் உள்ள தொகுதியிலிருந்து, ஒவ்வொரு கூறும் $\frac{m}{M}$ என்ற சமவாய்ப்புடன் அமையுமாறு தெரிந் தெடுக்கலாம்.

$n_i = n'$ என்று கொண்டால், ஒவ்வொரு துணைக்கூறும் $\frac{n'}{N'}$ என்ற சமவாய்ப்புடன் துணைமாதிரியில் இடம்பெறும்.

$\therefore$ மொத்தம் $mn' = n$ அளவுள்ள மாதிரியில் ஒவ்வொரு $(N_i = N')$ துணைக்கூறும் தெரிந்தெடுக்கப்படும் வாய்ப்புநிலை

$$\frac{m}{M} \cdot \frac{n'}{N'} = \frac{m}{M} \cdot \frac{n/m}{N/M} = \frac{n}{N}$$ என்றாகிறது.

எனவே, சுய எடையிட்ட மாதிரிகளில், $\bar{Y}$-ன் மதிப்பீடானது

$$\hat{\bar{Y}} = \frac{1}{mn'} \cdot \sum_{m} \sum_{n'} y_{ij} \qquad \ldots\ldots \quad (1)$$

என்று பெறப்படுகிறது. இதில் நிபந்தனைகளாவன :

$$N_i = \frac{N}{M} = N' \quad ; \quad n_i = \frac{n}{m} = n'.$$

அதாவது, திரள்களின் அளவும், துணைக்கூறுகளின் அளவும் N', n' என்று முறையே அமைதல் வேண்டும்.

இதில், (i) ஒரு துணைக்கூறும் மாதிரியில் தெரிவு பெற $\frac{n}{N}$ என்ற வாய்ப்பு உள்ளதாகவும் (ii) மாதிரிச் சராசரியானது தொகுதிச் சராசரியின் எளிய (எடையிடாத) மதிப்பீடாகவும் அமைதல் குறிப்பிடத்தக்கதாகும்.

எனவே, எடையிடாத மாதிரிச் சராசரியானது, முழுமைத் தொகுதிச் சராசரியின் பிறழ்ச்சியற்ற மதிப்பீடாக அமைவதை நிறுவலாம்.

இதன்மூலம், திரள் அளவு (cluster size) மாறுபடினும், மாதிரிச் சராசரியானது சிறந்த மதிப்பீடாக அமையக் காணலாம். அதாவது, முதல் கட்டமாகிய திரள்களை மாதிரியிடுகையில், திரளின் அளவு கருதப்படும்.

திரள்களினின்று மாதிரி எடுக்கையில் (முதல் கட்டம்) தெரிந்தெடுக்கப்படும் ஒவ்வொரு திரளும் வெவ்வேறு (மாறு

படும்) வாய்ப்பினைப் பெறும். அதாவது $\frac{N_i}{N}$ என்றபடி வாய்ப்பு மாறுபடும்.

உதாரணமாக, நான்கு தொகுதிகளில் அல்லது திரள்களில் ($M = 4$) உள்ள மதிப்புகள் Y_1, Y_2, Y_3, Y_4 என்று கொண்டால்,
$Y = Y_1 + Y_2 + Y_3 + Y_4$ என்றாகும்.

Y மதிப்பை, $m=1$ என்ற அளவில் திரளினை எடுத்து (மாதிரி) அதன்மூலம் மதிப்பிடுவோம். Y_2 என்ற மதிப்பைத் தெரிவு செய்வதாகக் கொண்டால், $\hat{P}_2$ என்பதை மொத்த மதிப்பில் Y_2-வின் வீத மதிப்பீடாகக் கருதி,

$$\hat{Y} = \frac{Y_2}{\hat{p}_2}$$ எனப் பெறலாம்.

மேலும்,

$$\hat{p}_2 = \frac{Y_2}{Y}$$ என்று கொள்ள,

$$\hat{Y} = \frac{Y_2}{\hat{p}_2} = \frac{Y_2}{Y_2/Y} = Y$$

என்று கண்டு, Y-ன் முற்றிலும் துல்லிய மதிப்பீடு $\frac{Y_2}{\hat{p}_2}$ என்று பெறப்படும். பொதுவாக,

$$\hat{Y}_2 = \frac{Y_i}{\hat{p}_i}$$ ஆகும்.

ஆனால், மாதிரி எடுக்கையில், p-மதிப்புகள் அறியப்படாத நிலையில், பின்வருமாறு கருதுகிறோம்.

Y_i-ன் வாய்ப்பு மதிப்புகள் (probabilities) தெரிந்திருக்கலாம். இவற்றை p_i என்று குறிப்பிடுக.

$$\therefore \quad \hat{Y} = \frac{Y_i}{p_i}$$ என்றாகும்.

(p_i என்பதை, $\frac{Y_i}{Y}$ மதிப்பின் தோராய மதிப்பாகக் கொள்க.)

பொதுவாக, $\hat{p}_i$ என்ற வாய்ப்பு அளவினை, $\hat{p}_i$-க்குப் பதிலாகப் பயன்படுத்துகையில், $\hat{Y}$ ஆனது Y-ன் பிறழ்ச்சியற்ற மதிப்பீடாக,

$$E(\hat{Y}) = \sum_{M} p_i \frac{Y_i}{p_i} = \sum_{M} Y_i = Y$$

என்று அறியப்படும்.

$m = 2$ எனும் அளவில் திரள்கள் மாதிரியானது அமையப் பெறின்,

Y_2, Y_3 என்ற மதிப்புகளைக் கருதி,

$\frac{Y_2}{p_2}$, $\frac{Y_3}{p_3}$ என்னும் மதிப்பீடுகள், Y-க்குப் பெறப்படுகின்றன.

அதாவது,

$$\hat{Y} = \frac{1}{2}\left(\frac{Y_2}{p_2} + \frac{Y_3}{p_3}\right)$$

$\{ E(\hat{Y}) =$ என்று காட்டுக $\}$.

p_i என்ற வாய்ப்பு மதிப்புகளை அறிய PPS முறையான முதல்நிலை மாதிரித் தெரிவில் பயன்படுத்தப்படுகிறது.

அதாவது, N_i என்ற அளவில் அமையும் முதல்நிலைக் கூறுகளுக்கேற்ப, வீத அளவில் (அல்லது திரள்களின் அளவு விகிதப்படி), முதல்கட்ட மாதிரியானது அமைகிறதெனலாம்.

$\frac{N_i}{N}$ என்றபடி அமையும் வாய்ப்புகளை (p_i மதிப்புகள்), $\frac{Y_i}{Y}$ — க்குப் பதிலாகப் பயன்படுத்தலாம். எப்படி ?

திரள்களின் உட்கூறுகள் பலபடித்தானதாக (Heterogeneous) அமைந்தால் அத் திரள்களின் இடைப்பட்ட வேறுபாடு (between clusters) குறைவதை, ஏற்கெனவே அறிந்தோம்.

3 திரள்களை ($m = 3$)

	y_{ij}	Y_i	$\bar{Y}_i$
$N_1 = 2$	3, 7	10	5
$N_2 = 5$	1, 4, 6, 11, 3	25	5
$N_3 = 3$	2, 5, 8	15	5

சராசரிகளிடை வேறுபாடில்லை ($\bar{Y}_i = 5$).

ஆனால், திரள் உள்ளீட்டு வேறுபாடு அதிகமாக உள்ளது.

விதங்கள் :

$$\frac{Y_1}{Y} = \frac{10}{50} = \frac{1}{5}; \quad \frac{Y_2}{Y} = \frac{25}{50}; \quad \frac{Y_3}{Y} = \frac{15}{50}.$$

மேலும், திரள் அளவுக்கேற்ற வாய்ப்பு விதங்கள்

$$\frac{N_1}{N} = \frac{2}{10} = \frac{1}{5}; \quad \frac{N_2}{N} = \frac{5}{10}; \quad \frac{N_3}{N} = \frac{3}{10}.$$

$$\therefore \quad \frac{N_i}{N} = \frac{Y_i}{Y} \text{ ஆகிறது.}$$

இனி, இரண்டாம் கட்டத் தெரிவில், $m = 2$ என்ற திரள்களினின்று, n_1, n_2 அளவுள்ள துணைமாதிரிகள் (subsamples) எடுக்கலாம்.

இவற்றின் மூலம், Y_1, Y_2 இவற்றை மதிப்பிட்டு,

$\hat{Y}_1$, $\hat{Y}_2$ என்று குறிப்போம்.

திரும்பவைக்காத முறையில் எடுக்கப்படும் துணைமாதிரிகளின் மொத்தங்கள் (totals),

$$y_1 = \sum_{n_1} y_{ij} \quad ; \quad y_2 = \sum_{n_2} y_{2j} \text{ ஆகும்.}$$

SRS முறைப்படி,

$$\hat{Y}_1 = \frac{N_1}{n_1}. \; y_1 \; ; \quad \hat{Y}_2 = \frac{N_2}{n_2}. \; y_2 \text{ என்று வரும்.}$$

பொதுவாக,

$$\hat{Y}_i = \frac{N_i}{n_i} \cdot y_i \text{ என்பதாகும்.} \qquad \ldots\ldots (2)$$

மேலும்,

$E(\hat{Y}_i) = Y_i$ என்பது முடிவு.

இரு கட்டங்களையும் இணைத்து, (2)-ஐ,

$$\hat{Y} = \frac{1}{m} \sum_m \frac{Y_i}{p_i} \text{ என்பதில் பிரதியிட,}$$

$$\hat{Y} = \frac{1}{m} \sum_m \frac{\hat{Y}_i}{p_i}$$

$$= \frac{1}{m} \sum_m \frac{1}{p_i} \cdot \left(\frac{N_i}{n_i} \cdot y_i \right)$$

$$\therefore \quad \hat{Y} = \frac{1}{m} \sum_m \frac{1}{p_i} \cdot \left(\frac{N_i}{n_i} \sum_{n_i} y_{ij} \right) \qquad \ldots (3)$$

என்று பெறப்படும்.

இதில் p_i என்பது, Y_i என்ற மதிப்புத் தெரிவு செய்யப்படும் வாய்ப்பாகும்.

$$\text{மேலும் } E(\hat{Y}) = E_i \left(E_j \frac{1}{m} \sum_m \frac{1}{p_i} \hat{Y}_i \right)$$

$$= E_i \frac{1}{m} \sum_m \left(\frac{1}{p_i} \cdot E_j(\hat{Y}_i) \right)$$

$$= \frac{1}{m} \sum_m \sum_M p_i \left(\frac{1}{p_i} \cdot Y_i \right)$$

$$\therefore \quad E(\hat{Y}) = \sum_M Y_i = Y$$

$\therefore$ $\hat{Y}$ என்ற (3)-ஆனது Y-ன் பிறழ்ச்சியற்ற மதிப்பீடாகக் காட்டப்பட்டது.

பயிற்சி : $p_i = \frac{1}{M}$ என்று பிரதியிட்டுப் பெறப்படும் $\hat{Y}$ ஆனது, எளிய திரள் மாதிரி முறையில் அமைவதெனக் காட்டு.

PPS-மாதிரி முறையில், $p_i = \frac{N_i}{N}$ என்று பிரதியிட, (3) ஆனது,

$$\hat{Y} = \frac{1}{m} \sum_{m} \frac{N}{N_i} \cdot \frac{N_i}{n_i} \cdot y_i$$

$$= \frac{N}{m} \sum_{m} \frac{y_i}{n_i}$$

$$\therefore \quad \hat{Y} = \frac{N}{m} \sum_{m} \frac{1}{n_i} \sum_{n_i} y_{ij} \qquad \dots (4)$$

என்றாகும்.

துணை முடிவு :

$n_i = \frac{n}{m} = n'$ எனில், (4) ஆனது

$\hat{Y} = \frac{N}{mn'} \sum_{m} \sum_{n'} y_{ij}$ என்றாகும்.

$\therefore$ முழுத் தொகுதிச் சராசரியின் மதிப்பீடு,

$\hat{\bar{Y}} = \frac{1}{mn'} \sum_{m} \sum_{n'} y_{ij}$ என்று பெறப்படும்.

இது, n-அளவுள்ள மாதிரியின் சராசரியாகும்.

இதனை

$$\bar{Y} = \bar{y}_{pps} = \frac{1}{mn'} \sum_{m} \sum_{n'} y_{ij}$$

என்று வரையறுக்கிறோம் ($n = mn'$).

'*pps*' என்ற அடிக்குறியானது, மாதிரி முறையைக் காட்டும் குறியீடாகும்.

$\bar{y}_{pps}$-ன் மாறுபாடு (Sampling Variance of $\bar{y}_{pps}$)

வரையறைப்படி, $\bar{y}_{pps}$-ன் மாதிரிசார் மாறுபாடானது

$$V(\bar{y}_{pps}) = E(\bar{y}_{pps} - \bar{Y})^2$$

என்பதாகும்.

$$\therefore \quad E(\bar{y}_{pps}) = \bar{Y}.$$

எதிர்பார்க்கும் மதிப்பைக் காணுதற்கு,

$$(\bar{y}_{pps} - \bar{Y}) = \left(\bar{y}_{pps} - \frac{1}{m}\sum_{m}\bar{Y}_i\right) + \left(\frac{1}{m}\sum_{m}\bar{Y}_i - \bar{Y}\right)$$

$$= u + V$$

என்பதனைக் கருதுக.

இதன்மூலம்

$$V(\bar{y}_{pps}) = E(\bar{y}_{pps} - \bar{Y})^2$$
$$= E(u + V)^2$$
$$= E(u^2) + E(2u \cdot V) + E(V^2)$$

என்று அமையும்.

முதலில், *i*-ஆவது திரள் நிலையானதெனக் கொண்டு, *j*-மூலம் எதிர்பார்க்கும் மதிப்பைப் பெறுகிறோம். அதன்பின் பல்வேறு *i* மதிப்புகளின் அடிப்படையில், எதிர்பார்க்கும் மதிப்பினைக் காண, $V(\bar{y}_{pps})$ பெறப்படும். இதன்படி :

$$2uV = 2\left(\frac{1}{m}\sum_{m}\bar{y}_i - \frac{1}{m}\sum_{m}\bar{Y}_i\right)\left(\frac{1}{m}\sum_{m}\bar{Y}_i - \bar{Y}\right)$$

$$= 2 \cdot \frac{1}{m}\sum_{m}(\bar{y}_i - \bar{Y}_i) \cdot \left(\frac{1}{m} \cdot \sum_{m}\bar{Y}_i - \bar{Y}\right)$$

$$\therefore \quad E_i\,[E_j\,(2uV)]$$

$$= E_i\left[E_j \cdot 2 \cdot \frac{1}{m}\sum_{m}(\bar{y}_i - \bar{Y}_i)\left(\frac{1}{m}\sum_{m}\bar{Y}_i - \bar{Y}\right)\right]$$

u, V என்பவை ஒன்றோடொன்று தொடர்பற்றவையாகவும்,
($\because$ இறுதித் திரள்கள் தொடர்பற்ற நிலையிலும் உள்ளன.)

மேலும் $E(\bar{y}_i) = \bar{Y}_i$ என்றும் அமைவதால்,

$$E_j(2uV) = 0 \text{ என்று ஆகிறது.}$$

அடுத்து

$$u^2 = \left(\frac{1}{m}\sum_m \bar{y}_i - \frac{1}{m}\sum_m \bar{Y}_i\right)^2$$

$$= \frac{1}{m^2}\left[\sum_m (\bar{y}_i - \bar{Y}_i)\right]^2$$

$$= \frac{1}{m^2}\left[\sum_m (\bar{y}_i - \bar{Y}_i)^2 + 2\sum_m \sum_{i \neq i'} (\bar{y}_i - \bar{Y}_i)(\bar{y}_{i'} - \bar{Y}_i)\right]$$

$$\therefore E_i[(E_j(u^2)] = E_i\left[E_j \cdot \frac{1}{m^2}\sum (\bar{y}_i - \bar{Y}_i)^2 + E_j \cdot 2\sum_i \sum_{i \neq i'} (\bar{y}_i - \bar{Y}_i)(\bar{y}_i - \bar{Y}_i)\right]$$

இறுதித் திரள்கள் தொடர்பற்றவை ஆதலால், வலப்புற இரண்டாம் உறுப்பு பூச்சியமாகும்.

மேலும் $E_i\left[E_j \cdot \frac{1}{m^2}\sum (\bar{y}_i - \bar{Y}_i)^2\right]$

$$= E_i \cdot \left(\frac{1}{m^2} \cdot \sum_m \frac{N_i - n'}{N_i} \cdot \frac{S_i^2}{n'}\right)$$

என்றாகிறது. இதில்

$$S_i^2 = \frac{1}{N_i - 1}\sum_{N_i} (y_{ij} - \bar{Y}_i)^2.$$

எனவே,

$$E_i\,[E_j\,(u^2)] = \frac{1}{m^2}\cdot m\cdot \sum_{M} \frac{N_i}{N}\cdot \frac{N_i - n'}{N_i}\cdot \frac{S^2}{n'}$$

$$= \frac{1}{mN}\sum_{M} (N_i - n')\cdot \frac{S_i^2}{n'}$$

ஆகிறது. ஏனெனில், $i : 1, 2, \ldots, M$ எனவும், i-ஆவது முதல் நிலைக்கூறினைத் தெரிந்தெடுக்கும் வாய்ப்பு $\frac{N_i}{N}$ என்றும் கருதப்பட்டது.

எனவே, $E_i\,(u^2) = \frac{1}{mN}\sum_{M}(N_i - n')\cdot \frac{S^2}{n'}$ ஆகும்.

இறுதியாக, $$V^2 = \left(\frac{1}{m}\sum_{m} \bar{Y}_i - \bar{Y}\right)^2$$

$$= \left[\frac{1}{m}\sum_{m} (\bar{Y}_i - \bar{Y})\right]^2$$

$$= \frac{1}{m^2}\cdot\left[\sum_{m} (\bar{Y}_i - \bar{Y})^2 + 2\sum_{i}\sum_{i\neq i'} (\bar{Y}_i - \bar{Y})\;(\bar{Y}_{i'} - \bar{Y})\right]$$

முதல் நிலை மாதிரிக் கூறுகள் திரும்ப வைக்கப்படும் முறையில் தெரிந்தெடுக்கப்படுவதாலும், $E_i(\hat{Y}_i) = \sum_{M}\frac{N_i}{N}\cdot\bar{Y}_i = \bar{Y}$ என்பதாலும் வலப்புற 2 ஆம் உறுப்பு பூச்சியமாகிறது.

$$\therefore\quad E_i\,(V^2) = \frac{1}{m^2}\cdot E_i\left[\sum_{m} (\bar{Y}_i - \bar{Y})^2\right]$$

$$= \frac{1}{m^2}\cdot m\sum_{M}\frac{N_i}{N}\;(\bar{Y}_i - \bar{Y})^2$$

$$\therefore \quad E\,(V^2) = \frac{1}{m.N} \sum_{M} N_i\,(\bar{Y}_i - \bar{Y})^2$$

இம் மூன்று முடிவுகளையும் பயன்படுத்தி,

$$V\,(\bar{y}_{pps}) = \frac{1}{mN} \sum_{M} N_i\,(\bar{Y}_i - \bar{Y})^2 + \frac{1}{mN} \sum_{M} (N_i - n')\frac{S_i^2}{n'}$$

என்றாகிறது.

எனவே, $\bar{y}_{pps}$-ன் மாதிரிமுறை சார்ந்த மாறுபாடாவது,

$$V\,(\bar{y}_{pps}) = \frac{1}{mN} \sum_{M} N_i\,(\bar{Y}_i - \bar{Y})^2 + \frac{1}{mN} \sum_{M} (N_i - n').\ \frac{S_i^2}{n'} \quad \ldots \quad (5)$$

இதில்,

$$S_i^2 = \frac{1}{(N_i - 1)} \cdot \sum_{N_i} (y_{ij} - \bar{Y}_i)^2.$$

இம் மதிப்பீட்டில் வலப்புறம் முதல் உறுப்பானது திரள்களின் சராசரிகளின் இடைப்பட்ட மாறுபாடாகவும், இரண்டாம் உறுப்பானது திரள்களின் உட்கூறுகளின் உள்ளீட்டு மாறு பாடாகவும் கருதப்படும்.

[$S_i^2 = i$ என்ற திரளில் உள்ள y_{ij}-ன் மாறுபாட்டு அளவாகும்.]

துணைமுடிவு 1

$N_i = n'$ என்று அமைந்திடின், அதாவது, திரள் முழுதும் மாதிரி எனக் கருதப்படும் நிலையில், முதல் கட்டத்தினை மட்டும் கொண்டு

$$V\,(\bar{y}_{pps}) = \frac{1}{mN} \cdot \sum_{M} N_i\,(\bar{Y}_i - \bar{Y})^2 \quad \ldots \quad (6)$$

என்று மதிப்பிடலாம். அதாவது, $\hat{Y}, Y$ இவற்றின் இடையில் உள்ள வேறுபாடு சிறிய அளவினதானால், $V\,(\bar{y}_{pps})$ ஆனது குறையும் தன்மையுடையதாகும்.

துணைமுடிவு 2

இரண்டாம் உறுப்பின் பகுதி அளவு (denominator) $m.n'$ ஆவதால், m, n' இவற்றின் பங்கீடு, இவை இம் மதிப்பை அவ்வளவாகப் பாதிப்பதில்லை. ஆனால், முதல் உறுப்பின் பகுதி அளவாக m அமைவதால், m-ஐப் பெரிதாகவும், n'-ஐச் சிறிதாகவும் கொண்டு, $V\,(\bar{y}_{pps})$-ன் அளவினை மேலும் குறைக்கலாம்.

$V\,(\bar{y}_{pps})$-ன் பிறழ்ச்சியற்ற மாதிரி அளவை மதிப்பிடுதல் :

$\hat{V}\,(\bar{y}_{pps})$-ன் மதிப்பானது எளிதாக அமைகிறது.

அதாவது,

$$\hat{V}\,(\bar{y}_{pps}) = \frac{1}{m\,(m-1)} \sum_{m} (\bar{y}_i - \bar{y}_{pps})^2 \qquad \ldots \quad (7)$$

இதில்,

$$\bar{y}_i = \frac{y_i}{n_i} \quad ; \quad \bar{y}_{pps} = \frac{1}{mn'} \sum_{m} \sum_{n'} y_{ij}$$

$$= \frac{1}{m} \sum_{m} \bar{y}_i$$

(7)-ல் கண்டபடி $\hat{V}(\bar{y}_{pps})$ ஆனது பின்வருமாறு நிறுவப்படும்.

தேற்றம் :

$$\hat{V}\,(\bar{y}_{pps}) = \frac{1}{m\,(m-1)} \cdot \sum_{m} (\bar{y}_i - \bar{y}_{pps})^2$$

என்பது $V\,(\bar{y}_{pps})$-ன் பிறழ்ச்சியற்ற மதிப்பீடாகும்.

நிரூபணம் :

$E\,[\hat{V}\,(\bar{y}_{pps})] = V\,(\bar{y}_{pps})$ என்று காட்டவேண்டும்.

$$(\bar{y}_i - \bar{y}_{pps}) = \bar{y}_i - \frac{1}{m} \Sigma \bar{y}_i$$

$$= \bar{y}_i - \frac{1}{m} . \ \bar{y}_i - \frac{1}{m} \sum_{i' \neq i} \bar{y}_{i'}$$

$$\therefore \ (\bar{y}_i - \bar{y}_{pps}) = \frac{m-1}{m} . \ \bar{y}_i - \frac{1}{m} \sum_{i' \neq i} \bar{y}_{i'}$$

என்பதன்மூலமாக.

$$\hat{V} \ (\bar{y}_{pps})$$

$$= \frac{1}{m} \cdot \frac{1}{(m-1)} \cdot \sum_{m} \left(\frac{m-1}{m} \cdot \bar{y}_i - \frac{1}{m} \sum_{i' \neq i} \bar{y}_{i'} \right)^2$$

$$= \frac{1}{m} \left[\frac{1}{m-1} \sum_{m} \left\{ \left(\frac{m-1}{m} \cdot \bar{y}_i \right)^2 \right. \right.$$

$$\left. \left. + \left(\frac{1}{m} \sum_{i' \neq i} \bar{y}_{i'} \right)^2 - 2 \left(\frac{m-1}{m} \cdot \bar{y}_i \right) \left(\frac{1}{m} \sum_{i' \neq i} \bar{y}_{i'} \right) \right\} \right]$$

அதாவது $\hat{V} \ (\bar{y}_{pps}) = \frac{1}{m}[u + v + w]$ ஆகிறது.

இதில்,

$$u = \frac{1}{(m-1)} \sum_{m} \left(\frac{m-1}{m} \cdot \bar{y}_i \right)^2$$

$$= \frac{m-1}{m^2} \cdot \sum_{m} \bar{y}_i^2.$$

$$v = \frac{1}{m-1} \sum_{m} \left(\frac{1}{m} \sum_{i' \neq i} \bar{y}_{i'} \right)^2$$

$$= \frac{1}{m^2 (m-1)} \cdot \sum_{m} \left(\sum_{i' \neq i} \bar{y}_{i'} \right)^2$$

$$= \frac{1}{m^2 (m-1)} \sum_{m} \left(\sum_{i:1}^{m} \bar{y}_i - \bar{y}_i \right)^2$$

$$\therefore \quad V = \frac{1}{m^2 (m-1)} \cdot \sum_{m} \left[\left(\sum_{m} \bar{y}_i \right)^2 - 2 \bar{y}_i \sum_{i} \bar{y}_i + \bar{y}_i^2 \right]$$

$$= \frac{1}{m^2 (m-1)} \left[m \left(\sum_{m} \bar{y}_i \right)^2 - 2 \left(\sum \bar{y}_i \right)^2 + \sum_{m} \bar{y}_i^2 \right]$$

$$= \frac{1}{m^2 (m-1)} [(m-2)(\Sigma \bar{y}_i)^2 + \Sigma \bar{y}_i^2]$$

$$= \frac{1}{m^2 (m-1)} \cdot \left[(m-2) \left(\sum_{m} \bar{y}_i^2 + \sum_{i' \neq i} \bar{y}_i \cdot \bar{y}_{i'} \right) + \Sigma \bar{y}_i^2 \right]$$

$$= \frac{1}{m^2 (m-1)} \cdot \left[(m-1) \sum_{m} \bar{y}_i^2 + (m-2) \sum_{i \neq i'} \bar{y}_i \cdot \bar{y}_{i'} \right]$$

$$= \frac{1}{m^2} \sum_{m} \bar{\bar{y}}_i^2 + \frac{m-2}{m^2 (m-1)} \sum_{i \neq i'} \bar{\bar{y}}_i \cdot \bar{\bar{y}}_{i'}$$

மேலும்

$$w = \frac{1}{(m-1)} \sum_{m} \left[2 \left(\frac{m-1}{m} \cdot \bar{\bar{y}}_i \right) \cdot \left(\frac{1}{m} \sum_{i \neq i'} \bar{\bar{y}}_{i'} \right) \right]$$

$$= \frac{2}{m^2} \sum_{m} \bar{\bar{y}}_i \cdot \sum_{i \neq i'} \bar{\bar{y}}_{i'}.$$

$$\therefore \quad v = \frac{2}{m^2} \sum_{i \neq i'} \bar{y}_i \cdot \bar{y}_{i'}.$$

இவற்றைப் பிரதியிட்டு,

$$\hat{V}(\bar{\bar{y}}_{pps}) = \frac{1}{m}\Bigg[\frac{m-1}{m^2} \sum_{m} \bar{y}_i^2 + \frac{1}{m^2} \sum_{m} \bar{\bar{y}}_i^2$$

$$+ \frac{m-2}{m^2(m-1)} \sum_{i \neq i'} \bar{\bar{y}}_i \cdot \bar{\bar{y}}_{i'} - \frac{2}{m^2} \sum_{i \neq i'} \bar{y}_i \cdot \bar{y}_{i'} \Bigg]$$

$$= \frac{1}{m}\Bigg[\frac{1}{m} \sum_{m} \bar{\bar{y}}_i^2 - \frac{1}{m(m-1)} \sum_{i \neq i'} \bar{\bar{y}}_i \cdot \bar{\bar{y}}_{i'} \Bigg]$$

அதாவது

$$\hat{V}(\bar{\bar{y}}_{pps}) = \frac{1}{m} \cdot (A)$$

இனி, எதிர்பார்க்கும் மதிப்பினை, அதாவது E என்பதனை மதிப்பிடவேண்டும். முதலில், i-ஆவது திரளானது கொடுக்கப் பட்டதெனக் கொண்டு, j-யின் மூலமாக (கட்டுப்பாட்டு எதிர் நோக்கும் மதிப்பு) E-யினை மதிப்பிட்டு, அடுத்ததாக, $i : 1, 2 \ldots M$ என்பதனைக் கருதி இறுதி மதிப்பினைப் பெற வேண்டும்.

$$\text{அதாவது } E(\hat{V}_{pps}) = E_i \left(E_j \cdot \frac{1}{m} \cdot A \right)$$

$$= \frac{1}{m} \cdot E_i (E_j \cdot A)$$

இதில்,

$$E_j\ (\mathrm{A}) = E_j \left[\frac{1}{m} \sum_{m} \bar{y}_i^2 - \frac{1}{m(m-1)} \sum_{i \neq i'} \bar{y}_i . \bar{y}_{i'} \right]$$

$$= E_j [\, B - C \,]$$

எனில்,

$$E_j\ (B) = E_j\left(\frac{1}{m}\sum_{m}\bar{y}_i^2\right) - \frac{1}{m}\sum_{m} E_j\,(\bar{y}_i^2)$$

$$= \frac{1}{m}\sum_{m}\left(\bar{\bar{y}}_i^2 + \frac{N_i - n'}{N_i}\cdot\frac{S_i^2}{n'}\right)$$

$$\left(\because\ E\bar{y}^2 = \bar{Y}^2 + E\,(\bar{y} - \bar{Y})^2 = \bar{Y}^2 + \frac{N-n}{N}\cdot\frac{S^2}{n}\right)$$

இதில்,

$$S_i^2 = \frac{1}{N_i - 1}\sum_{N_i}(y_{ij} - \bar{Y})^2,$$

$$E_j\,(C) = \frac{1}{m\,(m-1)}\,E_j\sum_{i \neq i'}\bar{\bar{y}}_i\cdot\bar{\bar{y}}_{i'}$$

$$= \frac{1}{m\,(m-1)}\sum_{i \neq i'}\bar{\bar{Y}}_i\cdot\bar{\bar{Y}}_{i'}$$

($\because$ இரண்டாம் கட்ட துணைமாதிரியானது SRS முறையில அமைகிறது. அதே i-ஆவது திரளானது மறுபடி தெரிந்தெடுக்கப்படின், அம் மாதிரியானது திரும்பவைக்கப்பட்டு மறுபடி, தொடர்பற்ற முறையில் i-ஆவது திரளினின்று எடுக்கப் பட்டதால், $\bar{y}_i$, $\bar{y}_{i'}$ ஆகியவை தொடர்பற்றிருக்கும்.)

எனவே,

$$E_i\,(E_j\cdot A) = E_i\,(E_j\cdot B - E_j\cdot C)$$

$$= E_i\left[\frac{1}{m}\sum_{m}\bar{Y}_i^2 + \frac{1}{m}\sum_{m}\frac{N_i - n'}{N_i}\cdot\frac{S_i^2}{n'} - \frac{1}{m\,(m-1)}\sum_{i \neq i'}\bar{Y}_i\cdot\bar{Y}_{i'}\right]$$

$$= E_i\,[\,D + F - G\,]$$

$$E_i(D) = \frac{1}{m}.\ E_i \sum_{m} \bar{\bar{y}}_i^2 = E_i \cdot \bar{Y}_i^2$$

$$= \bar{Y}^2 + \frac{1}{N} \sum_{M} N_i (\bar{Y}_i - \bar{Y})^2$$

$$E_i(F) = E_i \left(\frac{1}{m} \sum_{m} \frac{N_i - n'}{N_i}.\ \frac{S_i^2}{n'} \right)$$

$$= E_i \frac{N_i - n'}{N_i}.\ \frac{S_i^2}{n'} = \sum_{M} \frac{N_i}{N}.\ \frac{N_i - n'}{N_i}.\ \frac{S_i^2}{n'}$$

$$E_i(G) = E_i \left(\frac{1}{m}.\ \sum_{m} \frac{N_i - n'}{N_i}.\ \frac{S_i^2}{n'} \right)$$

$$= \sum_{M} \frac{N_i}{N}.\ \frac{N_i - n'}{N_i}.\ \frac{S_i^2}{n'}.$$

$$\therefore\ E_i(E_j \cdot A) = \bar{Y}^2 + \frac{1}{N} \sum_{M} N_i (\bar{Y}_i - \bar{Y})^2$$

$$+ \sum_{M} \frac{N_i}{N}.\ \frac{N_i - n'}{N_i}.\ \frac{S_i^2}{n'} - \bar{Y}^2$$

$$= \frac{1}{N} \sum_{M} N_i (\bar{Y}_i - \bar{Y})^2$$

$$+ \frac{1}{N} \sum_{M} (N_i - n') \frac{S_i^2}{n'}.$$

இறுதியாக,

$$E[\hat{V}(\bar{\bar{y}}_{pps})] = \frac{1}{m} E_i(E_j D)$$

$$= \frac{1}{mN} \cdot \sum_M N_i(\bar{Y}_i - \bar{\bar{Y}})^2$$

$$+ \frac{1}{mN} \sum_M (N_i - n').\ \frac{S_i^2}{n'}.$$

$$\therefore\ E[\hat{V}(\bar{\bar{y}}_{pps}) = V(\bar{\bar{y}}_{pps})$$

என்று நிறுவப்படுகிறது.

குறிப்பு :

நடைமுறையில், கணக்கீட்டு எளிமை கருதி,

$$\hat{V}(\bar{\bar{y}}_{pps}) = \frac{1}{m(m-1)} \sum_m (\bar{y}_i - \bar{\bar{y}}_{pps})^2$$

$$= \frac{1}{m(m-1)} \left[\sum_m \bar{y}^2{}_i - m.\bar{\bar{y}}_{pps}{}^2 \right]$$

என்று எழுதப்படுகிறது.

இதில்,

$$m\,\bar{\bar{y}}_{pps}{}^2 = \frac{1}{m(n')^2} \cdot \left(\sum_m y_i\right)^2 \text{ ஆகும்.}$$

PSS மாதிரி முறைக்கும், சாதாரணத் திரள் மாதிரி முறைக்கும் ஒரு ஒப்பீடு (Comparison of PPS and Cluster Sampling Methods)

ஒரு முழுமைத் தொகுதியின் மிகச் சிறந்த பிரதிநிதியாக உள்ளதொரு மாதிரியினைத் தெரிவு செய்தலின்மூலம், சாதாரணத் திரள் முறையைவிட அதிக அளவு திட்பமும் நுட்பமும் உடைய மதிப்பீடுகளைப் பெற வழியேற்படுகிறது. சாதாரணத் திரள் மாதிரி முறையில் அல்லது எளிய திரள் முறையில், வெவ்வேறு முதல் கட்டக் கூறுகள் (திரள்கள்) எளிய சமவாய்ப்பு முறையின்

அடிப்படையில் தேர்ந்தெடுக்கப்படுகின்றன. ஆனால், PPS முறையில், இவ்வாறல்லாமல், திரள் அளவிற்கேற்ற விகித அடிப்படையில் வாய்ப்புக்கள் பெற்ற விதத்தில் (probability proportional to size) முதல் கட்டத் தெரிவு நடைபெறுகிறது. இரண்டாம் கட்டத்தில், துணைமாதிரிக் கூறுகள் இருமுறைகளிலும் எளிய சமவாய்ப்பு முறைப்படி எடுக்கப்படுகின்றன.

எனவே, முதல்கட்டத் தெரிவில், திரட்களைத் தேர்ந்தெடுக்கும் பல்வேறு வழிமுறைகளுக்கேற்றபடி திட்பத்தின் அளவானது (precision) வேறுபடும் தன்மையைதாகும்.

எனவே, முதல்கட்டக் கூறுகளை (Primary sampling units) மட்டுமே, கருத்திற் கொண்டு, $V(\bar{\bar{y}}_{pps})$ மற்றும் $V(\hat{\bar{\bar{Y}}}_{cl})$ இவற்றை ஒப்பிட்டு அறியவேண்டும்.

ஏற்கெனவே கண்டறிந்தபடி,

$$V(\bar{\bar{y}}_{pps}) = \frac{1}{m \cdot N} \sum_{M} N_i (\bar{Y}_i - \bar{Y})^2$$

$$+ \frac{1}{mN} \sum_{M} (N_i - n'). \frac{S_i^2}{n'}$$

என்றும்,

$$V(\hat{\bar{\bar{Y}}}_{cl}) = \frac{1}{N^2}\left[M^2 . \frac{M-m}{M} . \frac{S_b^2}{m} + \frac{M}{m}\sum_{M} N_i^2 . \frac{N_i - n_i}{N_i} . \frac{S_i^2}{n'}\right]$$

என்றும் அமைகின்றன.

$N_i = n_i$ என்று கொள்வதன்மூலம், துணை மாதிரிக் கூறுகள் மாறுபாட்டு அளவினைப் பாதிப்பது தவிர்க்கப்படுகின்றது. அந் நிலையில்,

$$V(\bar{\bar{y}}_{pps}) = \frac{1}{mN} \sum_{M} N_i (\bar{Y}_i - \bar{Y})^2$$

$$V(\hat{\bar{\bar{Y}}}_{cl}) = \frac{M^2}{N^2} . \frac{M-m}{M} . \frac{S_b^2}{m}$$

$$= \frac{1}{(N')^2} \cdot \frac{M-m}{M} \cdot \frac{1}{m} \cdot \frac{1}{(M-1)} \sum_{M} (Y_i - \bar{Y})^2$$

என்று பெறுகிறோம்.

மேலும் $\frac{M-m}{M} = 1$ எனவும் $(M-1) \sim M$ எனவும் கொள்வதன் மூலம்,

$$V(\hat{\bar{Y}}_{cl}) = \frac{1}{m} \cdot \frac{1}{M} \sum_{M} \left(\frac{Y_i}{N'} - \bar{Y} \right)^2$$

$$V(\hat{\bar{Y}}_{cl}) = \frac{1}{m \cdot M} \cdot \sum_{M} \frac{(N_i \cdot \bar{Y}_i - N' \bar{Y})^2}{(N')^2} \qquad \ldots(A)$$

என்று சுருக்கமாய் அமைகிறது.

$$\text{மேலும் } V(\bar{\bar{y}}_{pps}) = \frac{1}{m} \sum_{M} \frac{N_i}{N} \cdot (\bar{Y}_i - \bar{Y})^2 \qquad \ldots(B)$$

என்று எழுதப்படும்.

(*A*), (*B*) இவற்றை ஒப்பிடுவதன்மூலம் இரு முறைகளுக்கும் ஒரு ஒப்புமையானது பெறப்படும்.

திரள் அளவுகளான N_i மதிப்புக்கள் அதிக அளவில் வேறுபடின், *PPS* முறையானது பயன் தரும் என முன்பு குறிப்பிட்டோம். மேலும், திரள்களின் Y_i மதிப்புகள், N_i-அளவுகளின் விகிதத்தில் மாறுபடின், $\bar{\bar{y}}_{pps}$ மதிப்பீடானது $\bar{Y}$ மதிப்புக்கு மிக அண்மையில் இருக்குமென்றும் கண்டறிந்தோம். அதாவது, Y_i அளவானது N_i அளவுடன் வீதத் தொடர்பில் வேறுபடுவதாக அமையும் நிலையில் (proportionally), *PPS* முறையானது விரும்பப்படும்.

இவ்விவரங்களைக் கருத்திற் கொண்டு, $V(\bar{\bar{y}}_{pps})$, $V(\hat{\bar{Y}}_{cl})$ ஆகியவை ஒப்பிடப்படுகின்றன. இரண்டிற்குமுள்ள வேறுபாட்டினைப் பின்வருமாறு பெறுகிறோம் :

$$V(\hat{\bar{Y}}_{cl}) - V(\bar{\bar{y}}_{pps}) = \left[\frac{1}{m \cdot M} \frac{\sum_{M} (N_i \bar{Y}_i - N' \cdot \bar{Y})^2}{(N')^2} \right] - \left[\frac{1}{m} \sum_{M} \frac{N_i}{N} (\bar{Y}_i - \bar{Y})^2 \right]$$

$$\therefore \quad V(\hat{\bar{Y}}_{cl}) - V(\bar{\bar{y}}_{pps}) = \frac{1}{m} \frac{\sum_{M} (N_i - N') \cdot \bar{Y}_i^2}{NN'}$$

$$+ \frac{1}{m} \cdot \frac{1}{N} \cdot \sum_{M} (N_i - N') \cdot (\bar{Y}_i^2 - \bar{Y}^2) \quad ...(C)$$

துணை முடிவு (1) :

இதிலிருந்து, $N_i = \frac{N}{M} = N'$ என்று அமையும் நிலையில்,

$$V(\hat{\bar{Y}}_{cl}) - V(\bar{\bar{y}}_{pps}) = 0$$

என்று வருவதால், இரண்டு வழிமுறைகளும் ஒரே அளவு திட்பமுடையனவாக கருதப்படுகின்றன.

அந்நிலையில், $\frac{N_i}{N}$ என்ற 'தெரிவு செய்யப்படும்' வாய்ப்பானது, $\frac{N'}{N} = \frac{N}{MN} = \frac{1}{M}$ என்று அமைய, ஒவ்வொரு முதல்நிலைக் கூறும் அல்லது திரளும் (cluster), $\frac{1}{M}$ என்ற சம அளவு வாய்ப்புடன் மாதிரித் தேர்வில் இடம் பெறும் ; அதாவது, எளிய சம வாய்ப்பு முறையே காணப்படுகிறது.

எனவே, $V(\hat{\bar{Y}}_{cl}) = V(\bar{y}_{pps})$

துணை அளவு (2) :

N_i அளவுகள் அதிகமாக வேறுபட்டு இருப்பின், $V(\hat{\bar{Y}}_{cl}) - V(\bar{y}_{pps}) > 0$ என்று வருகின்ற நிலையில், PPS மாதிரித் தேர்வு முறையானது, எளிய திரள் மாதிரித் தேர்வு முறையினை விட அதிக அளவில் திட்பமுடையதென முடிவு செய்யப்படும்.

(C)-யில் வலப்புற முதல் உறுப்பாவது :

$$\frac{1}{m} \frac{\sum_{M} (N_i - N')^2 \cdot \bar{Y}^2}{NN'} > 0$$

என்று அமைந்திருக்க, இரண்டாம் உறுப்பான

$$\frac{1}{m \cdot N} \sum_{M} (N_i - N') \quad (\bar{Y}_i^2 - \bar{Y}^2)$$

என்பதனைப் பொறுத்தவரையில், பல்வேறு கருத்துக்களைக் கவனத்தில் கொள்ளவேண்டியுள்ளது.

முதலில், $V(\hat{\bar{Y}}_{cl})$ அல்லது $V(\bar{y}_{pps})$ அளவினைக் குறைத்திடும் பொருட்டு, (இரு கட்ட) திரள் மாதிரி முறையினில், திரள் சராசரிகளின் (cluster means) இடையிட்ட மாறுபாட்டின் அளவினைக் குறைவாக இருக்கும்படியாக எடுத்துக்கொள்கிறோம். எனவே, நடைமுறை நிபந்தனைகட்கேற்ப, *PPS*-முறை பயன் படுத்தப்படுகையில், $(\bar{Y}_i^2 - \bar{Y}^2)$ என்ற வேறுபாடானது, $\bar{Y}^2$ அல்லது $(\bar{Y}_i^2)$ மதிப்புடன் மட்டும் ஒப்பிடப்படுகையில், குறைவாக இருக்கும் என்று எதிர்பார்க்கப்படுகிறது. எனவே $(\bar{Y}_i^2 - \bar{Y}^2)$ என்ற காரணி குறையும் தன்மையதாக உணரப்படும்.

அடுத்தபடி, $(N_i - N')$ என்ற வேறுபாடானது, N_i என்னும் திரள் அளவு மிகச் சிறியதாக இருப்பின், குறைவானதாக அமைய, $(N_i - N') < 0$ என்று ஏற்படும். எனவே, 2-வது உறுப்பானது, மாறி மாறி வரும் $+$, $-$ குறிகளுடன் அமையும் உறுப்புகளின் கூட்டுத் தொகையாகும். ஒவ்வொரு உறுப்பும் சிறிதான அளவாக அமையும்.

இந்நிலையில், $V(\hat{\bar{Y}}_{cl}) - V(\bar{y}_{pps}) > 0$ என்பது நிகழ வாய்ப்பு உண்டு. அதாவது, *PPS* முறையில், *SRS* முறையுடன் ஒப்பிடுகையில், அதிக திட்பம் காணப்படலாம்.

எடுத்துக்காட்டு 1 :

பின்வரும் முழுமைத் தொகுதி கொடுக்கப்படுகிறது :

திரள்	y_{ij}
1	1, 2, 3
2	5, 5
3	9, 9, 10.

(i) $\bar{Y}$-ன் மதிப்பு :

$$\bar{Y} = \frac{1}{N} \cdot \sum_{m} \sum_{n'} y_{ij}$$

$M = 3 \, ; N = N_1 + N_2 + N_3 = 8.$

$$\therefore \quad \bar{Y} = \frac{1+2+3+5+5+9+9+10}{8} = \frac{44}{8}.$$

(ii) $m = 2$ என்ற அளவில் மாதிரித் திரள்களைத் தெரிவு செய்து, $n' = 2$ என்ற அளவில் துணைமாதிரிகள் தெரிவு செய்து, பெறப்படும் எல்லா வழி மாதிரிகளையும் பின்வருமாறு அமைக்கலாம் :

திரள்களினின்று, $m = 2$ என்ற அளவில், மொத்தம் 9 வழிகளில் மாதிரிகள் தெரிவு செய்யப்படும்.

11,	12,	13
21,	22,	23
31,	32,	33

இவற்றிலிருந்து, $n' = 2$ என்ற அளவில் எடுக்கப்படும் துணை மாதிரிகளைக் கீழ்க்கண்ட விதத்தில் அமைக்கலாம் :

(11), (13) என்ற திரள் சேர்வுகட்கு $n' = 2$ என்ற மாதிரிகளாவன :

I	I	$\Sigma\,\Sigma\, y_{ij}$
1, 2	1, 2	6
	1, 3	7
	2, 3	8
1, 3	1, 2	7
	1, 3	8
	2, 3	9
2, 3	1, 2	8
	1, 3	9
	2, 3	10
		72

1	3	$\Sigma\ \Sigma\ y_{tj}$
1, 2	9, 9	21
	9, 10	22
	9, 10	22
2, 3	9, 9	23
	9, 10	24
	9, 10	24
1, 3	9, 9	22
	9, 10	23
	9, 10	23
		204

(11) மற்றும் (13) எனும் திரள் சேர்விணைக் கருதி, $E(\bar{y}_{pps})$ ன் மதிப்பினைக் கணக்கிடுவோம்.

1-ம் திரள் தெரிவு செய்யப்பட வாய்ப்பு : $\frac{3}{8}\left(\frac{N_i}{N}\right)$ ஆகும்.

எனவே, திரும்ப வைக்கும் முறைப்படி, (11) என்ற இணையின் (pair) தெரிவு வாய்ப்பு : $\frac{3}{8} \times \frac{3}{8} = \frac{9}{64}$ ஆகும்.

இத் திரள்களின் இணைச் சேர்விலிருந்து, $n' = 2$ என்ற அளவில், $\binom{3}{2}\binom{3}{2} = 9$ துணைமாதிரிகள் பெறப்படும்.

∴ எதிர்பார்க்கும் மதிப்பு :

$$\frac{3}{8} \times \frac{3}{8}. \left[\frac{1}{9 \times 4}(6 + 7 + 8 + 7 + 8 + 9 + 8 + 9 + 10)\right]$$

$$= \frac{9}{64} \times \frac{72}{36} = \frac{9}{32} \text{ ஆகிறது.}$$

இதேபோல், (13) என்ற இணையில்,

$$\frac{3}{8} \times \frac{3}{8} \times \left[\frac{1}{9 \times 4}(21 + 22 + \ldots\ldots + 23)\right]$$

$$= \frac{9}{64} \cdot \frac{1}{36} \cdot 204 = \frac{51}{64}.$$

என்று எதிர்பார்க்கும் மதிப்பு அமையும்.

இந்த முறையில், 9 மாதிரித்திரள்கட்கும், அதாவது (11), (12), (13) ...... (33) போன்றவைகட்குக் கணக்கிட வேண்டும். இந்த எதிர்பார்க்கும் மதிப்புகளின் கூட்டுத் தொகையானது $E(\bar{y}_{pps})$-ஐத் தரும்.

திரள் இணைகளில்,

(1, 1)-ன்	எதிர்பார்க்கும் மதிப்பு		:	$\frac{9}{64} \cdot \frac{72}{36}$
(1, 3)-ன்	,,	,,	:	$\frac{9}{64} \cdot \frac{204}{36}$
(1, 2)-ன்	,,	,,	:	$\frac{3}{8} \cdot \frac{2}{8} \cdot \frac{42}{12}$
(2, 1)-ன்	,,	,,	:	$\frac{3}{8} \cdot \frac{2}{8} \cdot \frac{42}{12}$
(2, 2)-ன்	,,	,,	:	$\frac{4}{64} \cdot \frac{20}{4}$
(2, 3)-ன்	,,	,,	:	$\frac{6}{64} \cdot \frac{86}{12}$
(3, 1)-ன்	,,	,,	:	$\frac{9}{64} \cdot \frac{204}{36}$
(3, 2)-ன்	,,	,,	:	$\frac{6}{64} \cdot \frac{86}{12}$
(3, 3)-ன்	,,	,,	:	$\frac{9}{64} \cdot \frac{1}{36} \cdot 336$

இவற்றைக் கூட்டிப் பெறும் மதிப்பாவது

$$E(\bar{y}_{pps}) = \frac{9}{32} + \frac{51}{64} + \frac{42}{128} + \frac{42}{128} + \frac{5}{16} + \frac{43}{64} + \frac{51}{64} + \frac{43}{64} + \frac{336}{256}$$

$$= \frac{1408}{256} = \frac{44}{8} = 5{\cdot}5.$$

$$= \bar{Y}$$

∴ $\bar{y}_{pps}$ பிறழ்ச்சியற்ற மதிப்பீடாகக் காண்க.

எடுத்துக்காட்டு 2 :

$M = 3$ திரள்களில் முறையே 200, 150, 250 மாணவர்கள் காணப்படுகின்றனர். $m = 2$ என்றபடி திரள்களைத் தெரிவு செய்து, $n' = 5$ என்ற அளவில் துணைக் கூறுகளை எடுத்து, மாணவர்கள் ஒருநாளைக்குச் சராசரியாகப் பிடிக்கும் சிகரெட்டுகளின் அளவை மதிப்பிட வேண்டியுள்ளது. PPS முறையில் :

$m = 2$ என்று திரள்களை எடுக்க,

	N_i	ΣN_i
A	200	1 — 200
B	150	201 — 350
C	250	351 — 600

(A, B, C திரள்கள்).

ராண்டம் எண் பட்டியலிலிருந்து, இரண்டு மூன்று இலக்க எண்கள், 231, 055 என்று எடுக்க, அதன்படி, A, B திரள்கள் கருதப்படும்.

$n' = 50$ என்று A, B-யில் எடுக்கப்படும் துணை மாதிரி மதிப்புக்கள் : (கற்பனை) :

A	B
y_{1j}	y_{2j}
3	4
4	7
5	8
6	4
7	3
25	22

$$\therefore \quad \bar{y}_{pps} = \frac{1}{mn'} \sum_{m} \sum_{n'} y_{ij}$$

$$= \frac{1}{2 \times 5} (25 + 22)$$

$$= \frac{47}{10} = 4{\cdot}7 \simeq 5$$

சராசரியாக 5 சிகரெட்டுகள் புகைக்கப்படுவதாகக் கொள்ளலாம்.

எடுத்துக்காட்டு 3 :

$M = 20$ கல்லூரிகளில் சராசரி எடை கணக்கிட மேற்கொண்ட ஆய்வில், *PPS* முறைப்படி $m = 5$ என்றபடியாகக் கல்லூரிகள் தெரிவு செய்யப்பட்டு, $n' = 20$ மாணவர்கள் ஒவ்வொரு கல்லூரியிலும் கருதப்படின், பின்வரும் பட்டியலைப் பயன்படுத்தி, $\bar{y}_{pps}$ மதிப்பினைக் காணுதல் :

கல்லூரி	N_i	$\sum_{n'} y_{ij}$
1	600	3000
2	800	3200
3	500	2900
4	900	3100
5	700	3300

(எடை *lb*-அளவில்)

$$\bar{y}_{pps} = \frac{1}{mn'} \sum_{m} \sum_{n'} y_{ij}$$

$$= \frac{1}{5 \times 20} \sum_{1}^{5} y_{ij}$$

$$= \frac{1}{(5) \cdot (20)} [3000 + 3200 + 2900 + 3100 + 3300]$$

$$\bar{\bar{y}}_{pps} = \frac{1}{100} (15500) = 155$$

$\therefore$ சராசரி எடை 155 *lb* என்று கணக்கிடப்படும்.

எடுத்துக்காட்டு 4 :

முதல் எடுத்துக்காட்டில் கொடுக்கப்பட்ட விவரங்கட்கு, $V(\bar{y}_{pps})$ கணக்கிடல் :

திரள்	y_{ij}	N_i	Y_i	$\bar{Y}_i$
1	1, 2, 3	3	6	2
2	5, 5	2	10	5
3	9, 9, 10	3	28	9·3

வரையறைப்படி,

$$V(\bar{y}_{pps}) = \frac{1}{mN} \sum_{M} N_i (\bar{Y}_i - \bar{Y})^2 + \frac{1}{mN} \sum_{M} (N_i - n') \cdot \frac{S_i^2}{n'}$$

$$= \frac{1}{2 \times 8} \left\{ \frac{3(2-5\cdot4)^2 + 2(5-5\cdot4)^2 + 3(9\cdot3-5\cdot4)^2}{3} \right\} + \frac{1}{2 \times 8} \left[\frac{1}{2}(1 + 0 + 0\cdot34) \right]$$

இங்கு $\begin{bmatrix} m = 2 \\ n' = 2 \end{bmatrix}$

$$\left\{ S_1^2 = \frac{1}{(N_i - 1)} \sum_{N_i} (y_{ij} - \bar{Y}_1)^2 \right.$$

$$= \frac{(1-2)^2 + (2-2)^2 + (3-2)^2}{2}$$

$$= 1.$$

$$\left. S_2^2 = 0 ; \quad S_3^2 = \frac{1}{2}(0\cdot09 + 0\cdot09 + \cdot49) = 0\cdot34 \right\}.$$

எனவே,

$$V(\bar{y}_{pps}) = \frac{1}{48} [11\cdot56 \times 3) + (0\cdot16 \times 2) + (15\cdot21 \times 3)] + \frac{1}{16} \left[\frac{1\cdot34}{2} \right]$$

$$= \frac{80 \cdot 63}{48} + \frac{0 \cdot 67}{16} = \frac{82 \cdot 64}{48}$$

$\therefore \quad V(\bar{y}_{pps}) = 1 \cdot 72.$

எடுத்துக்காட்டு 5 :

$M = 5$; $m = 2$ என்றபடி மாதிரிகள் எடுக்கப்பட்டு, $n' = 5$ எனும்படி துணைக் கூறுகள் கருதப்படின், பின்வருமாறு அமையும் நிலையில், கொடுக்கப்பட்ட துணை மாதிரியினைக் கருதி $\hat{V}(\bar{\bar{y}}_{pps})$ மதிப்பைப் பெறுக :

திரள்	N_i
1	70
2	50
3	80
4	40
5	60

திரள் 1-ம், 3-ம் எடுக்கப்பட்டு, அவற்றின் துணைக் கூறுகள் $(n' = 5)$:

1	3
y_{1j}	y_{3j}
25	20
20	25
20	25
20	30
25	30

என்று தரப்படுகிறது.

$$\Sigma\, y_{1j} = 110\ ; \quad \Sigma\, y_{2j} = 130$$

$$\therefore \quad \bar{\bar{y}}_{pps} = \frac{y}{n} = \frac{240}{10} = 24.$$

$V(\bar{y}_{pps})$-ன் மாதிரி அளவை :

அதாவது

$$\hat{V}(\bar{\bar{y}}_{pps}) = \frac{1}{m(m-1)} \sum_{m} (\bar{y}_i - \bar{\bar{y}}_{pp}\,)^2$$

$$= \frac{1}{2(2-1)} [(22-24)^2 + (26-24)^2]$$

$$\left[\bar{\bar{y}}_i = \frac{y_i}{n'} \right]$$

$$\left\{ \text{மேலும்,}\ \bar{\bar{y}}_1 = \frac{110}{5} = 22 ; \quad \bar{\bar{y}}_2 = \frac{130}{5} = 26 \right\}$$

$$\therefore \ \hat{V} (\bar{\bar{y}}_{pps}) = \frac{1}{2} (4+4) = \frac{8}{2} = 4$$

ஆகிறது.

இதிலிருந்து $s (\bar{y}_{pps}) = \sqrt{4} = 2$ என்பது தரப்பிழையின் மாதிரி அளவையாக அறியப்படும்.

பயிற்சி

1. பின்வரும் கற்பனைத் தொகுதியானது கொடுக்கப் படுகிறது :

திரள்	Y_{ij}
A	1, 2, 3
B	5, 5
C	9, 9, 10

$\bar{Y}$ மதிப்பைக் காண்க.

$m = 2$; $n' = 2$ என்ற அளவில் அமையும் எல்லா வழி மாதிரிகளையும் தேர்ந்தெடுத்து, அவற்றின் $\bar{y}_{pps}$ மதிப்புகளைக் கணக்கிடுக.

$E (\bar{y}_{pps}) = \bar{Y}$ என்று அமைவதைக் காட்டுக.

2. ஒரு மாநிலமானது 15 (= M) மாவட்டங்களாகப் பிரிக்கப்பட்டு, ஒவ்வொரு மாவட்டத்திலும் N_i பண்ணைகள் இருப்பதாக அமைகின்றது. ஒவ்வொரு பண்ணையிலும் சராசரி யாக உள்ள கால்நடைகளின் எண்ணிக்கையை மதிப்பிட வேண்டியுள்ளது. $m = 5$ என்ற அளவில் முதல்நிலைக் கூறுகளை

PPS முறையில் தெரிந்தெடுத்து, $n' = 4$ என்ற அளவில் துணைக் கூறுகளையும் எடுத்து, பின்வருவனவற்றைக் கணக்கிடுக :

(i) $\bar{\bar{y}}_{pps}$ மற்றும் $\hat{Y}$ மதிப்புக்கள்.

(ii) $\hat{V}(\bar{\bar{y}}_{pps})$ மற்றும் $s(\bar{\bar{y}}_{pps})$ இவற்றின் மதிப்புக்கள்

(iii) $\bar{\bar{Y}}$-க்கு, 2σ நம்பக எல்லை அளவு.

கொடுக்கப்படும் விவரம் :

மாவட்டம் (திரள்)	N_i
1	9
2	13
3	11
4	16
5	7
6	14
7	13
8	6

மாவட்டம்	N_i
9	18
10	9
11	12
12	10
13	8
14	16
15	14

மாதிரிப் பண்ணைகளில் :

y_{1j}	22	21	23	18
y_{2j}	15	18	28	13
y_{3j}	11	15	21	27
y_{4j}	18	20	16	16
y_{5j}	22	22	24	12

3. 20 கல்லூரிகளில் (M), ஒவ்வொன்றிலும் N_i மாணவர் பயில்வதாக அறியப்படுகிறது. ஒவ்வொரு மாணவனின் சராசரி எடையினையும் நிர்ணயிக்கவேண்டியுள்ளது. $m = 5$ கல்லூரிகள் PPS முறையில் தெரிவு செய்யப்படுகின்றன. $n' = 20$ மாணவர்

கள் ஒவ்வொரு கல்லூரியிலிருந்தும் தெரிவு செய்யப்படுகின்றனர். பின்வரும் விவரம் கொடுக்கப்படுகிறது. $\bar{y}_{pps}$ மற்றும் $\hat{Y}$ ஆகியவற்றைக் காண்க :

கல்லூரி	N_i	$\sum_{n'} y_{ij}$
1	600	3000 (*lb*)
2	800	3200
3	500	2900
4	900	3100
5	700	3300

4. ஒரு விற்பனைக் கூடத்தின் அலுவலர்களின் சராசரி ஊதிய வீதத்தினை மதிப்பிட, மொத்தம் 15 விற்பனைக் கூடங்களினின்று $m = 4$ என்ற அளவில் PPS முறையில் மாதிரித் தேர்வு செய்து, ஒவ்வொரு விற்பனைக் கூடத்திலிருந்தும் 10 அலுவலர்களை (n') கருதி, பெறப்படும் விவரம் :

	(ரூ).	
y_{1j}	61 64 64 69 75	72 70 60 60 60
y_{2j}	66 76 76 63 66	73 60 61 65 68
y_{3j}	64 61 63 68 72	63 63 57 69 73
y_{4j}	84 78 58 71 66	75 69 62 77 61

$\bar{y}_{pps}$ மதிப்பினைக் கணக்கிட்டு, $\hat{V}(\bar{y}_{pps})$ என்ன என்றும் கணக்கிடுக.

PPS மாதிரிமுறையில் மாதிரிப் பங்கீடு (Allocation of Sample)

திரள் மாதிரிமுறையில், ' PPS ' வழிமுறையைப் பயன்படுத்துகையில், எடுக்கப்படவேண்டிய முதல் நிலைக் கூறுகளின் (m) அளவினையும், துணைக் கூறுகளின் அளவினையும் (n') தீர்மானம் செய்வது கருதத்தக்க ஒரு பிரச்சினையாக அமைகிறது. மற்ற முறைத் தேர்வு நிலைகளைப் போன்றே, இந் நிலையிலும் மதிப்பீட்டு மாறுபாட்டு அளவையான $V(\bar{y}_{pps})$ ஆனது மிகச் சிறிய (minimum); மதிப்பாக்கத் தக்க விதத்தில், m மற்றும் n' ஆகியவற்றின் மதிப்புகள் நிர்ணயிக்கப்படுகின்றன.

குடும்பங்களின் செலவுத் தன்மையை, குறிப்பிட்ட ஆய்வினில், கருதுகையில், எவ்வளவு நகர்களைக் கருதி (m) அவற்றினின்று எவ்வளவு குடும்பங்களை (n') தெரிவு செய்து ஆய்வு நடத்துவது என்பது, மிகவும் கவனிக்கப்படவேண்டிய பிரச்சினை ஆகும்.

$n_i = \frac{n}{m} = n'$ என்ற நிலையில், $\bar{\bar{y}}_{pps}$ என்ற மதிப்பீடானது, மாதிரிச் சராசரியின் மதிப்பாக அமைகிறது. இதற்கான மாதிரி சார்ந்த மாறுபாட்டு அளவு—அதாவது $\hat{V}\ (\bar{\bar{y}}_{pps})$மதிப்பானது—கணக்கிட எளிதாக அமையும் நிலையினையும் ஏற்கெனவே கண்டோம். எனவே, n' என்பதை ஒரே சீரானதாக (uniform) எல்லாத் திரள்களுக்கும் கருதுவது, நடைமுறை எளிமைக்கு உதவும்.

m முதல்நிலைக் கூறுகளை, PPS முறையைப் பயன்படுத்தித் தெரிவு செய்வதாகவும், இறுதிநிலைக் கூறுகள் எல்லாம் ஒரே அளவில் (n') திரும்பவைக்காத சமவாய்பு மாதிரிமுறைமூலம் தெரிவு செய்வதாயும் (SRSWOR) கொண்டு, பின்வருமாறு கருதுகிறோம் :

ஏற்கெனவே கொடுக்கப்பட்டபடி,

$$V\ (\bar{\bar{y}}_{pps}) = \frac{1}{m.N} \sum_{M} N_i\ (\bar{Y}_i - \bar{Y})^2$$

$$+ \frac{1}{mN} \sum_{i=1}^{M} (N_i - n').\ \frac{S_i^2}{n'}$$

என்பதாகும்.

மொத்தச் செலவுச் சார்பானது (cost function)

$$c = c_1.m + c_2.mn'$$

என்று கொடுக்கப்படுவதாகக் கொள்வோம்.

c_1 என்பது முதல்நிலைக்கூறு ஒவ்வொன்றையும் தெரிவு செய்ய ஆகும் சராசரிச் செலவு ;

c_2 என்பது இறுதிநிலைக்கூறு ஒவ்வொன்றையும் தெரிவு செய்ய ஆகும் சராசரிச் செலவு என்று அமைந்திருக்கும்.

எனவே, நமது பிரச்சினையாவது :

கொடுக்கப்பட்ட செலவில், மாறுபாடான $V(\bar{\bar{y}}_{pps})$ மதிப்பைச் சிறும அளவினதாக (minimum) ஆக்கத்தக்க m மற்றும் n' ஆகியவற்றைக் கண்டறிதல் ஆகும்.

கண்டறியும் முறை :

இலக்ராஞ்சின் பெருக்கும் காரணியை (Lagrangean multiplier) பயன்படுத்தி :—

மாறுபாட்டு அளவு

$$V(\bar{\bar{y}}_{pps}) = \frac{1}{m.N} \sum_{M} N_i (\bar{Y}_i - \bar{Y})^2$$

$$+ \frac{1}{mN} \sum_{M} (N_i - n'). \frac{S_i^2}{}$$

என்பதாகும்.

இதனை,

$$V(\bar{\bar{y}}_{pps}) = \frac{A_1}{m} + \frac{A_2}{mn'} - \frac{A_3}{m.N'}$$

என்று சுருக்கமாக எழுதுவோம். (இது இயல்பான முறை
இதில்,

$$A_1 = \frac{1}{N} \sum_{M} N_i (\bar{Y}_i - \bar{Y})^2 ;$$

$$A_2 = \frac{1}{N} \sum_{M} N_i.S_i^2$$

$$A_3 = \frac{1}{M} \sum_{M} S_i^2$$

என்று காணப்படும்.

செலவுச் சார்பானது,

$$c = c_1 m + c_2.mn'$$

என்பதனைப் பயன்படுத்தி, $V(\bar{\bar{y}}_{pps})$-ன் சிறும மதிப்பைக் காண வேண்டும். அதற்கு,

$$f = \frac{A_1}{m} + \frac{A_2}{mn'} - \frac{A_3}{mN'} + \lambda\ (c_1 m + c_2 mn' - c)$$

என்ற சார்பு அமைக்கப்படுகிறது.

வரையறைப்படி, நிபந்தனைகள் அமைகின்றன :

(i) $\frac{\partial f}{\partial m} = 0$ என்பதிலிருந்து,

$$-\frac{A_1}{m^2} - \frac{A_2}{m^2.n'} + \frac{A_3}{m^2.N'} + \lambda c_1 + \lambda c_2.n' = 0$$

(ii) $\frac{\partial f}{\partial n'} = 0$ என்பதிலிருந்து,

$$-\frac{A_2}{m\,(n')^2} + \lambda \cdot c_2 m = 0$$

எனவும் பெறப்பட்டு, இவற்றிலிருந்து,

$$\lambda = \frac{1}{c_1}\left(A_1 - \frac{A_3}{N'}\right),\ \text{மற்றும்}$$

$$\lambda = \frac{1}{c_2}.\left(\frac{A_2}{(n')^2}\right)$$

என்று இரு மதிப்புகள் கண்டறியப்படுகின்றன.

இவற்றைப் பயன்படுத்தி,

$$(n')^2 = \frac{c_1}{c_2}.\ \frac{A_2}{A_1 - (A_3/N')}$$

$$\therefore\quad (n')^2 = \frac{c_1}{c_2}.\ \frac{S_{wi}^2}{S_{wb}^2 - S_i'^2/N'}$$

$$[A_1 = S_{wb}^2\ ;\ A_2 = S_{wi}^2\ ;\ A_3 = (S_i'^2).]$$

என்றும்,

$$n' = \sqrt{\frac{c_1}{c_2}.\ \frac{S_{wi}^2}{S_{wb}^2 - (S_i'^2/N')}} \quad \ldots\ldots \text{(A)}$$

என்றும் பெறுகிறோம். இதிலிருந்து,

$$m = \frac{c}{c_1 + c_2 \cdot n'} \quad \text{......} \quad \textbf{(B)}$$

என்று எளிதாகக் காணலாம்.

S_{wb}^2 என்பது, திரள்களின் இடைப்பட்ட மாறுபாட்டைக் குறிப்பது (திரள் அளவுகளை, N_i என்று எடைகளாகக் கருதி). S_{wi}^2 என்பது, S_i^2 மதிப்புகளின் எடையிட்ட சராசரியாகும். $(S_i')^2$ என்பதும் S_i^2-ன் மற்றொரு சராசரியாகும்.

இவற்றிலிருந்து, **சில முடிவுகள் :**

A-யில் பெறப்பட்ட சமன்பாட்டின்படி, திரள்களின் இடைப் பட்ட மாறுபாடானது, (S_{wb}^2) அதிகமாக இருப்பின், n'-அளவைக் குறைத்து, m-ன் அளவை அதிகரிக்கவேண்டியதாகிறது. முதல் நிலைக் கூறுகளை அதிக அளவில் தெரிவு செய்து, துணை நிலைக் கூறுகளைக் குறைந்த அளவில் தெரிவு செய்வது பயன் தரும்.

அடுத்து, திரள்களின் உள்ளீட்டு மாறுபாடு (S_{wi}^2) சிறியதாக இருப்பின், n'-மதிப்பை சிறியதாக வைக்கவேண்டியிருக்கும். அதாவது, திரள்களில் இடம்பெறும் மாதிரிக் கூறுகள் ஓரினத் தன்மையுடன் விளங்கிடின், n'-அளவு சிறிதாகவும், m-அளவு பெரியதாகவும் கருதப்படலாம். இதேபோல் $S_i'^2$-ஐயும் கருதலாம்.

c_1, c_2 ஆகிய செலவுகளைப் பொறுத்த அளவில், முதல்நிலைக் கூறினை (primary unit) எடுக்கச் சராசரிச் செலவானது (c_1) அதிகமாக இருப்பின், அதிக அளவில் துணைநிலைக் கூறுகளைத் (secondary unit) தெரிவு செய்து, முதல்நிலைக் கூறுகளைக் குறைந்த அளவில் எடுக்கவேண்டுமென்றும், மாறாக, துணை நிலைக் கூறினைத் தெரிவு செய்ப ஆகும் சராசரிச் செலவு (c_2) அதிகமாக இருப்பின், அதிக அளவில் முதல்நிலைக் கூறுகளையும், குறைந்த அளவில் துணைநிலைக் கூறுகளையும் எடுக்கவேண்டு மென்றும் முடிவு செய்யப்படுதல் வேண்டும். இவ்வாறாக, n' அளவினைப்பற்றி ஒரு மதிப்பீடு தெரிய வருகிறது.

மேலும், m, n' ஆகியவற்றின் பங்கீடு பற்றிய ஒரு குறிப்பினையும் கருதலாம். $V(\bar{y}_{pps})$-ன் தன்மையை நோக்குகையில், மொத்த மாதிரி அளவு, $mn' = n$ கொடுக்கப்படும் நிலையில், அது n' மூலம் எந்த அளவிலும் பாதிக்கப்படுவதில்லை என்று உணரலாம். அடுத்து, மொத்தச் செலவினை எடுத்துக் கொண்டால், அதுவும் n' மூலம் பாதிக்கப்படாத நிலையைக்

($mn' = n$ என்று கொடுக்கப்பட) காணலாம். அதாவது, $V(\bar{\bar{y}}_{pps})$ மற்றும் மொத்தச் செலவு c ஆகியவற்றை பாதிக்கும் ஒரே காரணியாக 'm' என்ற அளவு அமைவதை அறியலாம். மேலும், m அதிகமாக நேரிடுகையில், $V(\bar{\bar{y}}_{pps})$ சிறியதாகவும், c அதிகரித்தும் இருக்கும் நிலை குறிப்பிடத்தக்கது. எனவே, m, n' ஆகியவற்றைப் பங்கிடும் பிரச்சினையானது, செலவு அதிகரிப்பிற்கும் திட்ப அளவின் குறைவிற்கும் இடையில் ஒருவித ஒத்துப்போகும் தன்மையினைக் குறித்ததாக அமைகிறது. இத்தகைய சமநிலையானது, (*A*), (*B*) ஆகியவற்றின்மூலம் கொடுக்கப்பட்ட n', m ஆகிய மதிப்புகள்மூலம் பெறப்படுகிறது.

அடுத்தபடியாக, S_{wb}^2 மற்றும் S_i^2 ஆகியவற்றின் மாறுபடும் இயல்பைக் குறித்து ஒரு விளக்கத்தைப் பெறலாம். அவை இரண்டுமே குறைந்த மதிப்பு உடையதாக இருப்பின், $V(\bar{\bar{y}}_{pps})$ ஆனது சிறியதாக இருக்க வாய்ப்பு உண்டு. ஒரே சமயத்தில் இவ்விரண்டு மாறுபாட்டு அளவுகளையும் சிறிதாக இருக்கும்படி அமைத்துக்கொள்வது விரும்பத்தக்கது. திரள்களின் இடைப் பட்ட மாறுபாடு (between clusters) (S_{wb}^2) சிறியதாக ஆக்கப் பட்டால், திரள் உள்ளீட்டு மாறுபாடான S_i^2 பெரிதாக அமைகிறது. இத்தகைய நிலையில், $\hat{Y}$ ஆனது Y-க்கு வெகு அண்மையில் அமையவும், அதன் விளைவாக $V(\hat{Y}_{pps})$ மிகச் சிறியதாக இருக்கவும் நேரிடும் [ஏற்கெனவே அறிந்தபடி]. அதாவது, S_{wb}^2 சிறியதாக இருப்பின், $\bar{\bar{y}}_{pps}$-ம் $\bar{\bar{Y}}$-க்கு அண்மையில் இருக்க, $V(\bar{\bar{y}}_{pps})$ ஆனது சிறிதாக இருக்கலாம். எனவே, S_{wb}^2 மதிப்பைச் சிறியதானதாகவும், S_i^2 மதிப்பை பெரியதாகவும் வைத்துக்கொள்வது முடிவுகளை அதிகமாக பாதிப்பதில்லை என்று குறிப்பிடலாம்.

உதாரணமாக, மக்கட்தொகை கணிப்புப்போன்ற பெரிய அளவில் அமைந்த ஆய்வுகளில், வெவ்வேறு கிராமங்களை முதல் நிலைக் கூறுகளாக உருவாக்கி, அவற்றை வெவ்வேறு படுகைகளாக அமைத்துக்கொண்டு, பின்பு ஒவ்வொரு படுகையிலிருந்தும் ஒரு முதல்நிலைக் கூறினைத் தெரிவு செய்வதைக் குறிப்பிடலாம். இதில், ஒவ்வொரு முதல்நிலைக் கூறும் அதிக அளவில் பல்வேறு இயல்புகள் கொண்டு, உள்ளீட்டு மாறுபாட்டுடன் காணப்படும். அதாவது, கிராமங்களில் விவசாயத் தொடர்புடைய மற்றும் அத்தொடர்பற்ற இயல்புகள் கருதப்

படலாம். முதல்நிலைக் கூறின் அளவு வரம்புடையதாக அமைக்கப்படுகிறது.

n'-உடன் ஒப்பிடுகையில், N_i ஆனது மிகவும் பெரிய அளவினதாக இருக்குமிடத்து, மாறுபாடானது

$$V(\bar{\bar{y}}_{pps}) = \frac{1}{mN}\sum_{M} N_i(\bar{\bar{Y}}_i - \bar{\bar{Y}})^2 + \frac{1}{mN}\sum_{M} N_i \cdot \frac{N_i - n'}{N_i} \cdot \frac{S_i^2}{n'}$$

$$= \frac{1}{mN}\sum_{M} N_i(\bar{\bar{Y}}_i - \bar{\bar{Y}})^2 + \frac{1}{mN}\sum_{M} \frac{N_i S_i^2}{n'}$$

என்றாகும். இதன்மூலம் n' ஆனது

$$n' = \sqrt{\frac{c_1}{c_2} \cdot \frac{S_{wi}^2}{S_{wb}^2}}$$

என்று எளிதாகக் கணக்கிடப்படும்.

{வலதுபுறத்தில், $-S_i'^2/N'$ என்ற உறுப்பினை நீக்கிவிட்டு}

நடைமுறையில், S_i^2, S_{wi}^2, S_{wb}^2, $S_i'^2$ ஆகிய மதிப்புகள் தெரியாத நிலையில் அவற்றின் மாதிரி மதிப்பீடுகளான

$$s_i^2 = \frac{1}{n_i - 1}\sum_{n_i}(y_{ij} - \bar{\bar{y}}_i)^2;$$

$$s_{wb}^2 = \frac{1}{m-1}\sum_{m}(\bar{\bar{y}}_i - \bar{\bar{y}}_{pps})^2;$$

$$s_{wi}^2 = \frac{1}{n}\sum_{m} n_i \cdot s_i^2$$

$$s_i'^2 = \frac{1}{m}\sum_{m} s_i^2$$

ஆகியவற்றினைப் பயன்படுத்தி

$$n' = \sqrt{\frac{c_1}{c_2} \cdot \frac{s_{wi}^2}{s_{wb}^2 - (s_i'^2/N')}}$$

என்றும்

$$m = \frac{c}{c_1 + c_2 n'}$$

என்றும் பெறலாம்.

பயிற்சி

$M = 15$ என்று கொடுக்கப்பட்டு, N_i மதிப்புகள் கொடுக்கப்படுகின்றன :

திரள்	N_i	திரள்	N_i	திரள்	N_i
1	8	6	6	11	6
2	12	7	5	12	8
3	4	8	12	13	4
4	20	9	5	14	9
5	14	10	24	15	13

$m = 6$ என்ற அளவில், *PPS* முறையில் முதல்நிலைக் கூறுகளைத் தெரிந்தெடுக்கவும்.

இவை ஒவ்வொன்றிலிருந்தும் $n' = 4$ என்ற அளவில் துணை நிலைக் கூறுகள் தெரிந்தெடுக்கப்படுகின்றன. உதாரணமாக, பின்வரும் மாதிரியினைக் கருதுக :

திரள் எண்	y_{ij}
5	2, 3, 2, 5
8	2, 5, 5, 4
10	1, 4, 8, 7
11	4, 8, 9, 7
12	1, 1, 6, 4
15	4, 3, 3, 6

$c_1 =$ ரூ. 10 என்றும், $c_2 =$ ரூ. 2 என்றும் முறையே சராசரி செலவுகளைக் கருதி, உத்தம அளவு m மற்றும் n' ஆகியவற்றை மதிப்பீடு செய்க.

9. இரு மதிப்பீடுகள்

I. விகித மதிப்பீடு :

நோக்கம்—R-ன் வரையறை—R-ன் மதிப்பீடு—விகித மதிப்பீட்டுப் பிறழ்ச்சி—V $(\hat{R})$ காணுதல்.—சமவாய்ப்பு மாதிரிமுறையில்...—படுகை மாதிரிமுறையில்...—இணைத்தமுறை—தனித்தமுறை,—ஒழுங்குமாதிரிமுறையில். திரள் மாதிரி முறையில்—பிறழ்ச்சியற்ற—விகிதமதிப்பீடு—ஒரு நிலை.

II. மாறிகளின் தொடர்பு மதிப்பீடு :

நேர்கோட்டுத் தொடர்பு — சமவாய்ப்பு மாதிரிமுறை — இயல்பான நிலையோடு ஒப்பீடு—படுகை மாதிரிமுறை—ஒப்பீடு.

1. விகித மதிப்பீடு
(Ratio Estimate)

நோக்கம் :

பலவகையான மாதிரித் தேர்வு முறைகளைக் கருதி, அவற்றின் மூலம் முழுமைத் தொகுதியின் கூட்டு மொத்தம், சராசரி, வீத அளவு போன்ற பல மதிப்புகளைப் பெறும் விதத்தைப் பலவாருக ஆராய்ந்து கண்டறிந்தோம். ஆனுல், பொதுவாக நடைமுறையில் தேவைப்படுவன ஏதேனும் இரு இயல்புகளின் மொத்த மதிப்புகளின் இடையில் உள்ள விகித மதிப்புக்களே ஆகும். சமூகப் பொருளாதார ஆய்வுகளில், அதிகமாகக் கருதப்படுபவை, ஒவ்வொரு நபருக்கும் தலா ஆகும் செலவினம் (per capita expenditure), ஃஃகு தலா வருமானம் (per capita income), ஆண்-பெண் விகிதம், வேலையற்ருேர் விகிதம் போன்ற பல வகை மதிப்பீடுகளே ஆகும். அதேபோல், விளைவு ஆய்வுகளில் (crop surveys), உற்பத்தி விகிதம், மூலப் பொருள்கள் விகிதம் போன்றவற்றின் மதிப்பீடுகளே பெரும்பாலும் தேவைப்படுகின்றன. இத்தகைய விகித அளவுகளை மதிப்பிடுகையில்,தொகுதி அளவு மற்றும் பகுதி அளவு ஆகியவற்றின் பிறழ்ச்சியற்ற மதிப்பீடுகள் பெறப்பட்டு, தக்கபடி பயன்படுத்தப்படுகின்றன. இத்தகைய மதிப்பீட்டு அளவையானது, விகித மதிப்பீட்டு அளவை(ratio estimator)என்று வரையறுக்கப்படுகிறது. மொத்த அளவுகள் (uuits) தெரிந்திராத பட்சத்தில், முழுத் தொகுதி

சராசரி மற்றும் வீத அளவு இவற்றை மதிப்பிட்டு அறியவும் இத்தகைய மதிப்பீட்டு அளவைகள் இன்றியமையாதனவாகும்.

அதாவது, ஏதேனும் ஒரு மாதிரித் தேர்வு முறையில், மாதிரிகள் தெரிந்தெடுக்கப்பட்டு அவற்றின் அடிப்படையில் 'r' என்ற விகித அளவு கணக்கிடப்பட்டு, அதன் மூலமாக முழுமைத் தொகுதி விகிதமான R என்பது மதிப்பீடு செய்யப் படுகிறது.

தொகுதி மற்றும் பகுதி (Numerator & denominator) அளவுகளின் மதிப்பீடுகள் வீதத் தொடர்புற்று இருப்பின், விகித மதிப்பீட்டுமுறை அதிகத் திறனுடன் அமையும். பகுதியானது கொடுக்கப்படுகையில், மதிப்பீடுகளின் விகிதத்தைப் பகுதியின் உண்மை மதிப்பால் பெருக்கி வரும் மதிப்பானது தொகுதி உறுப்பின் மதிப்பீட்டு அளவையாகக் கருதப்படும். சில நிபந்தனை களுடன், விகித மதிப்பீட்டு முறையானது திட்பத்தினை அதிகரிப்பதாய் அமைவதைப் பின்பு காண்போம்.

விகித அளவு R-ன் வரையறை :

Y, X இரண்டும் இரு மாறிகளான y, x இவற்றின் முழுமைத் தொகுதிக் கூட்டு மொத்தங்களெனில், விகித அளவானது

$$R = \frac{Y}{X} \qquad \ldots\ldots \quad (1)$$

என வரையறுக்கப்படும். ஏதேனும் ஒரு மாதிரிமுறையைப் பயன்படுத்தி இதனை மதிப்பிட வேண்டும்.

உதாரணமாக, நான்கு குடும்பங்களில் கருதப்படும் நபர்கள் விவரம் :

குடும்பம்	ஆண்கள் (Y)	மொத்த நபர்கள் (X)
1	3	5
2	4	7
3	3	4
4	5	9
	15	25

என்றிருப்பின்,

$$X = \Sigma X_i = 25$$
$$Y = \Sigma Y_i = 15$$

எனப் பெற்று, ஆண்கள் விகிதமானது

$$R = \frac{Y}{X} = \frac{15}{25} = 0{\cdot}6$$

என்று பெறப்படும். அதாவது, முழுமைத்தொகுதியில், மொத்த நபர்களில் 60 சதவீதம் (60%) ஆண்கள் என்று அறியப்படும்.

இதனை மதிப்பீடு செய்யும் முறை :

$\hat{Y}$, $\hat{X}$ என்பவற்றை, Y, X ஆகியவற்றின் பிறழ்ச்சியற்ற மாதிரி மதிப்பீடுகளாகக் கொண்டால், R என்ற விகிதத்தின் மதிப்பீடு

$$\hat{R} = \frac{\hat{Y}}{\hat{X}} \qquad \ldots\ldots \quad (2)$$

என்று காணப்படும்.

இதிலிருந்து, Y-ன் விகித மதிப்பீடானது

$$\hat{Y}_R = \hat{R}{\cdot}X = \frac{\hat{Y}}{\hat{X}}.X$$

என்று நிர்ணயிக்கப்படும். இதற்கு, X என்ற துணை மாறியினைப் பற்றிய விவரமானது அறியப்பட வேண்டியது அவசியமாகும்.

மேற்கண்ட உதாரணத்தில், $n = 2$ என்று மாதிரிகளைக் கருதி R-ன் மதிப்புப் பெறப்படுவதாகக் கொள்வோம். மொத்தம் $\binom{4}{2} = 6$ வழிகளில் இம் மாதிரிகளைத் தெரிவு செய்கிறோம். அவற்றிலிருந்து, மாதிரி விகிதங்கள் பெறப்படும்.

மாதிரி	குடும்பங்கள்	மாதிரி விகித மதிப்பு
1	1, 2	$r_1 = \frac{y_1 + y_2}{x_1 + x_2} = \frac{(3+4)}{(5+7)} = \frac{7}{12}$
2	1, 3	$r_2 = \frac{3+3}{5+4} = \frac{6}{9}$
3	1, 4	$r_3 = \frac{3+5}{5+9} = \frac{8}{14}$
4	2, 3	$r_4 = \frac{4+3}{7+4} = \frac{7}{11}$
5	2, 4	$r_5 = \frac{4+5}{7+9} = \frac{9}{\ }$
6	3, 4	$r_6 = \frac{3+5}{4+9} = \frac{8}{13}.$

எனவே, மாதிரி விகிதம் r என்பதை,

$$r = \frac{\Sigma y}{\Sigma x} = \frac{y}{x}$$

என்று குறிப்பிடலாம். நம் உதாரணத்தில், r என்பது, மாதிரியிலுள்ள ஆண்களின் வீதத்தைக் (proportion) காட்டுவதாகக் கொள்வோம்.

ஆனால், 'r' என்பது, பொதுவாக அறியப்படும் 'வீத அளவு' என்னும் கருத்தினின்று வேறுபட்ட தோன்றும். n என்பது மொத்த அளவாக அமைந்திடின், $\frac{\Sigma y_i}{n}$ என்பது 'p' என்ற வீதமாகக் கருதப்படும். இதில் 'பகுதி' யானது (அதாவது n-மதிப்பு) மாதிரி அளவாகும்; அதவாது இது ஒரு 'ராண்டம்மாறி' (random variable) அல்ல. மாறாக, தொகுதியானது (Σy_i) ஒவ்வொரு மாதிரியிலும் மாறுபடக்கூடிய ராண்டம் மாறியாக அமையும்.

இங்கு கருதப்படும் குடும்பங்களில், y, x இரண்டு மதிப்புகளும் மாதிரிக்கு மாதிரி வேறுபடுவனவாகக் காணலாம். எனவே, $\frac{y}{x}$ என்பது இரு ராண்டம் மாறிகளுக்கிடையில் உள்ள விகித அளவாகும். எனவே, புள்ளிவிவர நியதிப்படி, r என்பது 'p' என்ற கருத்தினின்று முற்றிலும் வேறுபடுகிறது.

n குடும்பங்களைக் கருதி,

$y = \Sigma y_i$ = மொத்த குழந்தைகள் எண்ணிக்கை

$x = \Sigma x_i$ = மொத்த குடும்பங்கள்

என்று அமைந்தால் $r = \frac{y}{x}$ என்பது, தலா ஒரு குடும்பத்தில் உள்ள குழந்தைகளைக் குறிக்கும். $r = 3$ எனில், ஒவ்வொரு குடும்பத்திலும் தலா மூன்று குழந்தைகள் இருப்பதாக அறியப்படும். இதேபோல் பல உதாரணங்களைக் கொள்ளலாம். ஏற்றுமதி விகிதம், விற்பனை விகிதம், சேமிப்பு விகிதம், செலவு விகிதம் போன்ற பலவகைகளில் இதனைக் கருதலாம்.

ஒரு விகிதத்தின் பகுதி அளவு (denominator) மாதிரி அளவு 'n' ஆக அமையாது, ஏதேனும் ஒரு வாய்ப்பு மாறியாக இருக்கும் நிலையிலும், மாதிரிக் கூறுகள் (sampling units), முதல்நிலைக் கூறுகளினின்று மாறுபடும் நிலையிலும், விகித மதிப்பீட்டு முறை

இன்றியமையாத ஒன்றாக விளங்குகிறது. மேலும், y, x இவற்றினிடையே ஒட்டுறவு (correlation) இருப்பின், $\bar{Y}$ அல்லது Y பற்றிய சிறந்த மதிப்பீட்டை இம்முறையானது அளிக்கிறது.

எடுத்துக்காட்டாக, மொத்த நெல் உற்பத்தியை ஒரு மாவட்டத்தில் மதிப்பிடவேண்டியதாகக் கொண்டு, 500 பண்ணைகளைக் கருதுவோம். இவற்றிலிருந்து 50 பண்ணைகளை மாதிரியாக எடுத்து, அவற்றில் மொத்த விளைச்சல் $y = \Sigma y_i = 1000$ (கலங்கள்) என மதிப்பிட்டால்,

$$\bar{y} = \frac{1000}{50} = 20 \text{ (கலம்)}$$

என்பது ஒரு பண்ணையின் சராசரி விளைச்சல் ஆகும் (மாதிரியில்). மொத்தம் $N = 500$ பண்ணைகள் இருப்பதால், மொத்த நெல் உற்பத்தியானது

$$\hat{Y} = N \, . \, \bar{y} = 20 \times 500 = 10000$$

என்று பெறப்படும். இது, முழுத் தொகுதி (popoulation)-யின் மொத்தமான Y உடன் ஒப்பிடுகையில் மிகக் குறைந்ததாகக் காணப்படும்.

மேலும், ஒரு ஏக்கரில் சராசரி உற்பத்தியானது எல்லாப் பண்ணைகட்கும் சமமானதாக அறிந்திடப்பெறின், அந்த விவரத்தையும் பயன்படுத்தலாம். மொத்த விளைச்சல் பரப்பு கொடுக்கப்பட்டு, அதை X என்று குறிப்பிட, ஒரு ஏக்கருக்கு சராசரியாக உற்பத்தியாகும் நெல் அளவு மதிப்பிடப்பட்டு, அதனை X-ஆல் பெருக்கலாம். $x_1, x_2 \ldots x_n$ என்பவை மாதிரியில் கருதப்பட்ட பண்ணைகளின் பரப்புகளாகக் கொண்டு,

$\Sigma x_i = x = 200$ ஏக்கர் எனில்,

$$r = \frac{y}{x} = \frac{1000}{200} = 5 \text{ (கலம்)}$$

என்பது R-ன் மதிப்பீடாக அமைகிறது.

$\therefore$ மொத்த நெல் உற்பத்தி அளவு

$$\hat{Y} = r \, . \, X = \left(\frac{y}{x}\right) . \, X = 5 \times 10000$$

$$= 50{,}000 \text{ (கலம்)}$$

என்று மதிப்பிடப்படுகிறது.

விகித மதிப்பீட்டின் பிறழ்ச்சி அல்லது ஒருபுறச் சாய்வு (BIAS)

பிறழ்ச்சியற்றதொரு மதிப்பீட்டினது மாறுபாடானது (Variance), பிறழ்ச்சி உடையதொரு மதிப்பீட்டின் சராசரி வர்க்கப் பிழை (MSE)-யினைவிட அதிகமாக இருப்பின், பிறழ்ச்சி உடைய மதிப்பீட்டினையே தெரிவு செய்கிறோம். நடைமுறையில் பயன்படுத்தப்படும் தெரிவு முறைகள் அனைத்திலும், விகித மதிப்பீடானது முழுமைத் தொகுதி விகிதத்தின் பிறழ்ச்சி உடைய மதிப்பீடாக அமையும் என்பது குறிப்பிடத்தக்கது.

உதாரணமாக, N அளவுள்ள தொகுதியிலிருந்து எளிய சமவாய்ப்பு முறைப்படி n அளவுள்ள மாதிரி தெரிவு செய்யப் பட்டு, $\bar{y}$, $\bar{x}$ என்றபடி மாதிரி மதிப்பீடுகள் பெறப்படின்,

$$\hat{R} = \frac{\bar{y}}{\bar{x}} \text{ என்னும் மதிப்பீடானது, } R = \frac{Y}{X}$$

என்பதன் பிறழ்ச்சியுடைய மதிப்பீடாகக் காணலாம்.

திரும்ப வைக்காத முறைப்படி, (SRSWOR),

$$E(\hat{R}) = \frac{1}{\binom{N}{n}} \cdot \sum_{1}^{\binom{N}{n}} \frac{\bar{y}_i}{\bar{x}_i}$$

என்ற எதிர்பார்க்கும் மதிப்பானது, சாதாரணமாக R-ன் மதிப்புக்கு சமமாக இருப்பதில்லை. ஆதலால் இது ஒருபுறச் சாய்வுள்ளதாக (biased) அமைகிறது.

மற்றொரு மதிப்பீடான

$$\hat{R}_1 = \frac{1}{n} \sum_{1}^{n} \frac{y_i}{x_i}$$ என்பதும் பிறழ்ச்சியுடைதாகிறது.

ஏனெனில்,

$$E(\hat{R}_1) = \frac{1}{N} \sum_{1}^{N} \frac{Y_i}{X_i} \neq R.$$

எனவே, எதிர்பார்க்கும் மதிப்பைக் காணவும், பிறழ்ச்சியின் அளவினை, விகித மதிப்பீட்டு அளவையான $\frac{\hat{Y}}{\hat{X}}$-க்கு வரையறுக் கவும், பின்வருமாறு கருதுகிறோம் :

$$e_1 = \frac{\hat{Y} - Y}{Y}\ ;\ e_2 = \frac{\hat{X} - X}{X}$$ என்று கொள்க.

$\hat{Y}$, $\hat{X}$ ஆகிய மதிப்பீடுகள் பிறழ்ச்சியற்றவை ஆதலால்,

$$E(e_1) = E(e_2) = 0\ ;\ V(e_1) = \frac{V(\hat{Y})}{Y^2}\ ;\ V(e_2) = \frac{V(\hat{X})}{X^2}$$

மேலும் $\text{Cov}\,(e_1, e_2) = \dfrac{\text{Cov}\,(\hat{X}, \hat{Y})}{XY}$.

$\hat{Y} = Y(1 + e_1)$, $\hat{X} = X(1 + e_2)$ என்று பிரதியிட,

$$\hat{R} = R \cdot \frac{(1 + e_1)}{(1 + e_2)}$$ என்று பெறப்படும். ...(3)

$$\therefore\ E(\hat{R}) = R \cdot E\left\{\frac{1 + e_1}{1 + e_2}\right\} = R \cdot E\,(1 + e_1)\,(1 + e_2)^{-1}$$

$|e_2| < 1$ ஆதலால், இரண்டாவது உறுப்பினை விரிவு படுத்தி எழுதி, ஒருங்கும் தன்மையதாகக் (convergent) காண லாம். இதன்மூலமாக, X-ன் எல்லா மதிப்பீடுகளும் $(\hat{X})$, $(0,\ 2X)$ என்னும் இடைவெளியில் அமைகின்றன. $\hat{X}$-ன் மாறுபாடானது குறைந்த அளவில் இருக்கவேண்டுமெனில், மாதிரி அளவானது (n) அதிகமாதல் வேண்டும்.

எனவே,

$$E(\hat{R}) = R \cdot E\,[(1 + e_1)(1 - e_2) + e_2^2 - \ldots)]$$
$$= R \cdot E\,[1 + (e_1 - e_2) + (e_2^2 - e_1 e_2) + \ldots]$$

e_1, e_2 மிகச் சிறியனவாக அமைவதால், அவற்றின் இரண்டும் அதற்கு மேலும் உள்ள அடுக்கு (power) எண்கள் வரும் உறுப்புகள் நீக்கப்படுகின்றன. எனவே,

$$E(\hat{R}) = R$$ என வரும்.

$\therefore$ $\hat{R}$ பிறழ்ச்சியற்றது. $[\because\ E(e_1 - e_2) = 0]$

ஆனால் 2-ம் அடுக்குக்கு மேற்பட்டவற்றை மட்டும் நீக்கி விடும் நிலையில், இத்தகைய பிறழ்ச்சியற்ற நிலை அமைவதில்லை. ஏனெனில், அந்நிலையில், உடன்தொடர்பும் கருதப்படுகிறது.

∴ பிறழ்ச்சியானது ($\hat{R}$-ஐச் சார்ந்து),

$$B(\hat{R}) = E(\hat{R} - R) = R\left[\frac{V(\hat{X})}{X^2} - \frac{\mathrm{Cov}(\hat{X}, \hat{Y})}{XY}\right]$$

$$\therefore \quad B(\hat{R}) = \frac{1}{X^2}\left[R \cdot V(\hat{X}) - \mathrm{Cov}(\hat{X}, \hat{Y})\right] \qquad \ldots(4)$$

B என்ற கெழு, பிறழ்ச்சியினைக் குறிக்கும்.

மாதிரியின் அளவு அதிகரிக்கப்பட, $B(\hat{R})$ குறைகிறது.

$B(\hat{R}) = 0$ என்பதிலிருந்து,

$$R = \frac{\mathrm{Cov}(\hat{X}, \hat{Y})}{V(\hat{X})}$$ என்ற மதிப்புப் பெறப்படுகிறது.

[இது, (0, 0) என்ற புள்ளியின் வழியாக அமையும். $\hat{X}$-ஐச் சார்ந்த $\hat{Y}$-ன் உடன் தொடர்புக் கோட்டினை (Regression line) நிர்ணயிக்கும்.]

எனவே, கொடுக்கப்பட்ட நிபந்தனைகளுக்கேற்ப, ஏற்படும் பிறழ்ச்சியினை (4)-ல் கண்டவாறு பெறுகிறோம். இது, பெரிய அளவினதான மாதிரிகளைக் கருதுகையில் அமைகிறது.

மேலே பெறப்பட்ட பிறழ்ச்சிகளின் மாதிரி மதிப்பீடானது

$$b(\hat{R}) = \frac{1}{\hat{X}^2}\left[\hat{R} \cdot v(\hat{X}) - \mathrm{Cov}(\hat{X}, \hat{Y})\right] \qquad \ldots(5)$$

என்று தரப்படும். இது $B(\hat{R})$-ன் பிறழ்ச்சியற்ற மதிப்பீடு அல்ல.

மிகப் பெரிய அளவில் எடுக்கப்படும் மாதிரிகளில், விகித மதிப்பீட்டின் பிறழ்ச்சியானது பொருட்படுத்தத் தக்கதல்ல.

$$\text{Cov}(\hat{R}, \hat{X}) = E(\hat{Y}) - E(\hat{R}) \, . \, E(\hat{X})$$

என்பதிலிருந்து,

$$B(\hat{R}) = -\frac{\text{Cov}(\hat{R}, \hat{X})}{X} = -\frac{1}{X} \cdot \rho(\hat{R}, \hat{X}) \cdot \sigma_{\hat{R}} \cdot \sigma_{\hat{X}}$$

என்று சரியான பிறழ்ச்சி பெறப்படுகிறது.

இதில், $\rho(\hat{R}, \hat{X}) = \hat{R}, \hat{X}$ இவற்றினிடையே உள்ள ஒட்டுறவு.

$\sigma_{\hat{R}}, \sigma_{\hat{X}}$ என்பவை தரப் பிழைகள்

எனவே, சார் பிழையானது (relative bias)

$$\left| \frac{B(\hat{R})}{R} \right| = \left| \rho(\hat{R}, \hat{X}) \right| . C(\hat{R}). \; C(\hat{X}) \qquad \ldots\ldots (6).$$

என்று காணப்படும். $C(\hat{R})$, $C(\hat{X})$ என்பவை முறையே $\hat{R}$, $\hat{X}$ இவற்றின், சராசரியினைச் சார்ந்த தரப்பிழைகள். இதிலிருந்து,

மாதிரி அளவானது அதிகரிக்கும்பொழுது, விகித மதிப்பீட்டின் பிறழ்ச்சியின் அளவு குறைகிறது என்று முடிவு செய்யப்படும்.
[$\because$ $C(\hat{R})$, $C(\hat{X})$ ஆகியவை சிறு அளவினதாகக் காணலாம்]
$\rho(\hat{R}, \hat{X})$-ம் ஏறக்குறைய பூஜ்யத்துக்கு அண்மையிலிருக்கும்.
($\hat{Y}$, $\hat{X}$ ஆகியவை விகிதத் தொடர்புடையவை).

$\hat{R}$-ன் பிழை வர்க்கச் சராசரி (Mean Square Error).

விகித மதிப்பீட்டு அளவையானது பிறழ்ச்சியுடையதாக இருப்பதால், வேறு ஏதேனும் மதிப்பீட்டு அளவையுடன் திறன் அளவு அடிப்படையில் அதனை ஒப்பிடும் பொருட்டு MSE கணக்கிடப்படுகிறது. வரையறைப்படி,

$\hat{R}$-ன் பிழைவர்க்கச் சராசரியானது,

$$M(\hat{R}) = E(\hat{R} - R)^2 \text{ என்று தரப்படும்.}$$

அதாவது,

$$M(\hat{R}) = R^2.E[(e_1 - e_2)^2(1 + e_2)^{-2}]$$
[(2)-ன் மூலம்]

$| e' | <$ என்றும்,

(e_1, e_2) இவற்றின் 2-க்கு மேற்பட்ட அடுக்கு எண்கள் வரும் (powers) உறுப்புகளைக் கருதாத நிலையிலும்,

$$M(\hat{R}) = R^2.E[(e_1 - e_2)^2]$$

$$= R^2.E\left[\frac{V(\hat{Y})}{Y^2} - 2\frac{\text{Cov}(\hat{X}, \hat{Y})}{XY} + \frac{V(\hat{X})}{X^2}\right]$$

$$\therefore \quad M(\hat{R}) = \frac{1}{X^2} \; V(\hat{Y}) - 2R.\text{Cov}(\hat{X}, \hat{Y}) + R^2 \cdot V(\hat{X})]$$

என்றாகும். ...... (7)

இதன் மதிப்பீடு,

$$m(\hat{R}) = \frac{1}{\hat{X}^2}[v(\hat{Y}) - 2\hat{R}\,\text{Cov}(\hat{X}, \hat{Y}) + \hat{R}^2 v(\hat{X})]$$

என்பதாகும்.

{ நடைமுறையில், இவற்றை $V(\hat{R})$, $v(\hat{R})$ என்று கொள்வோம் }.

விகிதமுறை மதிப்பீடு :

$\hat{Y}$, $\hat{X}$ ஆகியவை பிறழ்ச்சியற்ற மதிப்பீடுகளாக, X கொடுக்கப்பட்டிருக்க Y-மதிப்பு

$$\hat{Y}_R = \left(\frac{\hat{Y}}{\hat{X}}\right) \cdot X$$

என்று மதிப்பீடு செய்யப்படும். எனவே,

$$B(\hat{Y}_r) = \frac{1}{X}[R.V(\hat{X}) - \text{Cov}(\hat{X}, \hat{Y})]$$

என்று வரும்.

மேலும்,

$$V(\hat{Y}_R) = V(\hat{Y}) - 2R.\text{Cov}(\hat{X}, \hat{Y}) + R^2 V(\hat{X})$$

என்றும் பெறப்படும்.

[MSE க்குப் பதிலாக, V-யினை அதாவது மாறுபாட்டு அளவினைக் கருதுகிறோம்.]

இதிலிருந்து, $\hat{Y}$ — ஐவிட, $\hat{Y}_R$ ஆனது, பின்வரும் நிலைக்கேற்ப அதிக திறனுடையதாக அமையும்.

அதாவது, $2R\,\text{Cov}(\hat{X}, \hat{Y}) > R^2.V(\hat{X})$ என்றிருத்தல் வேண்டும்.

அல்லது, R நேர்க் கணியமாக $(+)$,

$$\rho\ (\hat{X}, \hat{Y}) > \frac{1}{2} \cdot \frac{C(\hat{X})}{C(\hat{Y})}$$

என்றிருக்கவேண்டும்.

மேற் காணப்பட்ட, பிழ்ச்சி மதிப்பு, மற்றும் மாறுபாட்டு அளவை ஆகியவற்றை, அடிப்படையான சில மாதிரிமுறைகளைக் கருதி, விகித மதிப்பீட்டு அளவைகட்கு நிர்ணயிக்கலாம். முதலில் வருவது எளிய சமவாய்ப்பு மாதிரித் தேர்வுமுறை ஆகும்.

(அ) எளிய ஏாண்டம் மாதிரிமுறையும் விகித மதிப்பீடும் :

$$R = \frac{Y}{X} = \frac{\frac{\sum Y_i}{N}}{\sum X_i}$$ என்ற முழுமைத் தொகுதியின்

மாதீரி அளவையானது,

$$\hat{R} = r = \frac{\frac{\sum y_i}{n}}{\sum x_i} = \frac{y}{x}$$ என்று மதிப்பிடப்படுகிறது.

இம் மதிப்பீடானது பிறழ்ச்சியுடைய, ஆனால் பொருத்தமான (consistent) மதிப்பீடாக அறியப்படும்.

எடுத்துக்காட்டு :

தங்குமிடங்களில் காலியாக உள்ள அறைகளின் விகித அளவைப் பெற, கருதப்பட்ட 10 இடங்களில் காணப்பட்ட நிலையாவது :

எண்	அறைகள் (x_i)	காலியாக உள்ளவை (y_i)
1	20	4
2	10	2
3	30	6
4	40	8
5	20	5
6	15	3
7	25	5
8	30	6
9	20	3
10	20	5

காலியாக உள்ள அறைகளின் விகித அளவு R-ன் மதிப்பீடு :

$$\hat{R} = r = \frac{\Sigma y_i}{\Sigma x_i} = \frac{47}{220} = 0{\cdot}21 \text{ ஆகும்.}$$

$\hat{R}$, அதாவது, r-ன் பிறழ்ச்சி மதிப்பு : (Bias of $\hat{R}$)

பொதுவாக பெறப்பட்ட பிறழ்ச்சியானது (bias), (4)-ன் படி,

$$B(\hat{R}) = E(\hat{R} - R)$$

$$= \frac{1}{X^2} [R.V(\hat{X}) - \text{Cov}(\hat{X}, \hat{Y})]$$

என்பதாகும்.

சமவாய்ப்பு மாதிரிமுறையின்மூலம்,

$$V(\hat{X}) = N^2 . \frac{S_x^2}{n} . \frac{(N - n)}{N} \text{ என்றும்,}$$

$$\text{Cov}\ (\hat{X},\ \hat{Y})\ N^2 \cdot \rho.\ \frac{S_x S_y}{n}.\ \frac{(N-n)}{N}$$

என்றும் பெற்று, அவற்றின் மூலம்

$$B(\hat{R}) = \frac{N-n}{Nn}.\ \frac{1}{\bar{X}^2}\ [\ R.S_x^2 - \rho.S_x.S_y\ ]$$

என்று காண்கிறோம்.

எனவே, $R.S_x^2 - \rho\ S_x.S_y = 0$

அல்லது $R.S_x = \rho.S_y$

என்று அமையும் நிலையில், பிறழ்ச்சியானது பூச்சியமாக, மதிப்பீடாது பிறழ்ச்சியற்று அல்லது ஒருபுறச் சாய்வு இன்றி இருக்கும். இந்நிலையானது, x-ன் மீது அமையும் y-மாறியின் தொடர்பானது (அல்லது பின்னடைவு) (Regression) (0, 0) என்ற புள்ளியின் வழியாகச் செல்லும் நேர்கோடாக இருப்பதைக் காட்டும்.

$$\therefore\quad E(\hat{R}) = R \text{ ஆகிறது.}$$

மேற்குறித்த பின்னடைவானது (0, 0) என்னும் புள்ளியில் (point) அமையாவிடின்,

$$\text{பிறழ்ச்சி} = B(\hat{R})$$

$$= \frac{N-n}{Nn} \cdot \frac{1}{(\bar{X})^2} \cdot (R\ .\ S_x^2 - \rho\ .\ S_x\ .\ S_y)$$

$$= \frac{N-n}{Nn} \cdot \frac{1}{(\bar{X})^2} \cdot \left[\frac{\bar{Y}}{\bar{X}} \cdot S_x^2 - \rho\ S_x\ S_y\right]$$

$$= \frac{N-n}{Nn} \cdot R\left[\frac{S_x^2}{(\bar{X})^2} - \rho \cdot \frac{\bar{X}}{\bar{Y}} \cdot \frac{S_x \cdot S_y}{(\bar{X})^2}\right]$$

$$\therefore\quad B(\hat{R}) = \frac{N-n}{Nn} \cdot R \cdot [C_x^2 - \rho \cdot C_x \cdot C_y]$$

[C என்ற குறியீடு மாறுபாட்டுக்கெழு (coefficient of variation) மதிப்புக்களைக் குறிக்கும்.]

C_x, C_y இவற்றின் மதிப்புகள் சிறியனவாவதலால், $|\rho| \leq 1$ என்று ஆகி, மாதிரி அளவான n-ன் மதிப்பு அதிகரிக்கப்படுகையில், பிறழ்ச்சியின் அளவு குறைய நேரிடலாம். இம் மதிப்பானது,

r என்ற மதிப்பீட்டினது தரப்பிழையினைவிடக் குறைவு என்று பின்பு நிரூபிப்போம்.

எடுத்துக்காட்டு : 1

கொடுக்கப்படும் முழுமைத் தொகுதியானது :

மாதிரிக் கூறுகள்	Y_i	X_i
1	1	2
2	2	1
3	3	4
4	4	3
5	5	5

(i) R-ன் மதிப்பாவது :

$$R = \frac{\Sigma Y_i}{\Sigma X_i} = \frac{15}{15} = 1.$$

(ii) $n = 3$ என்ற அளவுப்படி மாதிரிகளை, திரும்பவும் வைக்காத முறைப்படி தெரிவுசெய்து மாதிரி விகிதங்களைக் காணல் :

மாதிரி எண்	y_i	x_i	$r = \frac{\Sigma y_i}{\Sigma x_i}$
1	1, 2, 3	2, 1, 4	$\frac{6}{7} = 0{\cdot}86$
2	1, 2, 4	2, 1, 3	$\frac{7}{6} = 1{\cdot}17$
3	1, 2, 5	2, 1, 5	$\frac{8}{8} = 1{\cdot}00$
4	1, 3, 4	2, 4, 3	$\frac{8}{9} = 0{\cdot}89$
5	1, 3, 5	2, 4, 5	$\frac{9}{11} = 0{\cdot}82$
6	1, 4, 5	2, 3, 5	$\frac{10}{10} = 1{\cdot}00$

7	2, 3, 4	1, 4, 3	$\frac{9}{8} = 1{\cdot}13$
8	2, 3, 5	1, 4, 5	$\frac{10}{10} = 1{\cdot}00$
9	2, 4, 5	1, 3, 5	$\frac{11}{9} = 1{\cdot}22$
10	3, 4, 5	4, 3, 5	$\frac{12}{12} = 1{\cdot}00$

(iii) பிறழ்ச்சி அளவினை மதிப்பிடல் :

$B\ (R)$ என்னும் பிறழ்ச்சியானது

$$B\ (\hat{R}) = E\,(r - R)$$

$$= \frac{N-n}{N.n}\cdot\ \frac{1}{\bar{X}^2}\,[R.S_x{}^2 - \rho\ .\ S_xS_y]$$

என்று சமவாய்ப்பு முறையில் வரையறுக்கப்படும்.

$$\bar{X} = \frac{15}{5} = 3\ \ ;\ \ \bar{Y} = \frac{15}{5} = 3.\ \ ;\ \ \therefore\ R = \frac{\bar{Y}}{\bar{X}} = 1.$$

$$S_x{}^2 = \frac{10}{4} = 2{\cdot}5\ ;\ S_y{}^2 = 2{\cdot}5$$

$$\therefore\ S_x = 1{\cdot}58\ \ ;\ \therefore\ S_y = 1{\cdot}58$$

$$S_{xy} = \frac{8}{4} = 2{\cdot}0. \qquad \therefore\ \rho = \frac{2{\cdot}0}{(1{\cdot}58)\,(1{\cdot}58)} = 0{\cdot}8$$

எனவே, $\rho\ S_xS_y = 2{\cdot}0$

$$R.S_x{}^2 = 1 \times 2{\cdot}5 = 2{\cdot}5.$$

எனவே,

$$B\ (\hat{R}) = B\,(r)$$

$$= \frac{(5-3)}{5 \times 3}\cdot\ \frac{1}{3^2}\,[\,2{\cdot}5 - 2{\cdot}0]$$

$$\therefore\ B\,(r) = \frac{2}{15}\cdot\ \frac{1}{9}\,[0{\cdot}5] = \frac{1}{135} = 0{\cdot}007$$

என்று மதிப்பிடப்படுகிறது.

r ன் மாதிரிமுறை சார்ந்த மாறுபாடு : (Variance of ' r ')

அதாவது $V(r) = E(r-R)^2$-ன் மதிப்பை அறிய வேண்டும். N-ன் மதிப்பு சிறிதாக அமைந்திடின், $V(r)$-ஐக் கணக்கிடுவது எளிதாகிறது. மாறாக, N அதிக அளவினதாக இருப்பின், $V(r)$-ன் மதிப்பினை ஒரு மாதிரியின் அடிப்படையில் மதிப்பீடு செய்து பெறவேண்டியிருக்கிறது.

குறிப்பு : $M(\hat{R})$ என்று முன்பு (7)-ல் பெறப்பட்ட மதிப்பீடானது,

$$V(\hat{R}) = V(r) = \frac{1}{\bar{X}^2}\cdot [V(\hat{Y}) - 2R\cdot \text{Cov}(\hat{X}, \hat{Y}) + R^2.V(\hat{X})]$$

என்று கொள்ளப்படும். }

சமவாய்ப்பு முறை முடிவுகளைப் பயன்படுத்த,

$$V(\hat{R}) = V(r) = \frac{1}{\bar{X}^2}\left[N^2.\ \frac{N-n}{Nn}.\ S_y^2 - 2R\cdot\rho\,\frac{S_x\cdot S_y}{n}.\ N^2\ \frac{(N-n)}{N}. + R^2.N^2.\ \frac{N-n}{Nn}.\ S_x^2\right]$$

$$\therefore\ V(r) = \frac{(N-n)}{Nn(\bar{X})^2}[S_y^2 - 2\rho.R.S_xS_y + R^2.S_x^2]$$

என்று சுருக்கமான அளவில் பெறப்படும்.

இதே மதிப்பினை, அடிப்படைக் கருத்தினின்று, $(Y_i - RX_i)$ என்ற வேறுபாட்டைக் கருதுவதன்மூலமாக,

$$V(r) = \frac{1}{\bar{X}^2}.\ \frac{N-n}{Nn}.\ \sum_1^N \frac{(y_i - Rx_i)^2}{N-1}$$

என்ற விதியைப் பயன்படுத்தி $\sum_1^N (y_i - Rx_i)^2$ மதிப்பைக் கண்டு,

$$V(r) = \frac{1}{\bar{X}^2}.\ \frac{N-n}{Nn}(S_y^2 - 2\rho\cdot RS_xS_y + R^2S_x^2)$$

என்று அதே மதிப்பினைப் பெறலாம்.

நடைமுறையில் மதிப்பீட்டுத் திட்பத்தினைக் கருதி, $V(r)$ ஆனது மிகச் சிறியதாக இருப்பது விரும்பப்படுகிறது. $V(r)$-ன் வரையறைமூலம், $\rho = 1$ என்று அமைந்திடும் நிலையில், $V(r)$ மதிப்பானது மீச் சிறு (minimum) அளவினதாகிறது. அதாவது, x, y இருமாறிகளும் மிகவும் உச்ச அளவில் (perfect) நேர்க்கணிய உடன்தொடர்பு பெற்றிருப்பின், விகித முறையானது (ratio method) சிறும மாறுபாட்டு அளவினைப் பெறுவதாக விளங்கும்.

மாதிரி அளவு n-ஆனது, எந்த வகையில் பிறழ்ச்சித் தன்மையையும், $\hat{R}$-ன் மாதிரிசார்ந்த மாறுபாட்டு அளவினையும் பாதிக்க வல்லது என்பதை இப்பொழுது ஆராய்வோம் :

'ராண்டம்' முறையில்,

$$B(\hat{R}) = \frac{N-1}{N} \cdot \frac{1}{n} \cdot \frac{1}{\bar{X}^2} (RS_x^2 - \rho\, S_x S_y)$$

$$\& \; V(R) = \frac{N-n}{N} \cdot \frac{1}{n} \cdot \frac{1}{\bar{X}^2} (S_y^2 - 2\rho S_x S_y + R^2 S_x^2)$$

எனக் கண்டறிந்தோம்.

$$\therefore \quad \text{தரப்பிழை} = \sqrt{V(\hat{R})} \text{ ஆகிறது.}$$

மாதிரி அளவானது ஐந்து மடங்காக அதிகரித்து, $5n$ என்று அமையும் நிலையில் (உதாரணமாக), பிறழ்ச்சியானது $\frac{1}{5}$ பகுதி (அதாவது 20%) குறையவும், தரப்பிழையானது $\frac{1}{\sqrt{5}}$ அளவு குறையவும் வழியேற்படுகிறது. எனவே, பெரும்பாலும் அதிக அளவுடைய மாதிரிகளில், பிறழ்ச்சியானது பொருட்படுத்தத் தக்கதல்ல.

எடுத்துக்காட்டு 2 :

எடுத்துக்காட்டு (1)-ல் கருதப்படும் மாதிரிகளின் அடிப்படையில், $V(\hat{R})$ அதாவது $V(r)$ மதிப்பைக் கணக்கிடுகிறோம் :

மாதிரி எண்	$r = \frac{\Sigma y_i}{\Sigma x_i}$	$(r-R) = r-1$	$(r-R)^2$
1	0·86	− 0·14	0·0196
2	1·17	0·17	0·0289
3	1·00	0	0
4	0·89	− 0·11	0·0121
5	0·82	− 0·18	0·0324
6	1·00	0	0
7	1·13	0·13	0·0169
8	1·00	0	0
9	1·22	0·22	0·0484
10	1·00	0	0
			0·1583

அடிப்படை வரையறையாவது :

$$V(r) = E(r-R)^2$$

$$\therefore \quad V(r) = \frac{\Sigma (r-R)^2}{10} = \frac{0{\cdot}1583}{10} = 0{\cdot}01583$$

($R = 1$ என்று கணக்கிடப்பட்டது).

இது, மாதிரிகளின் அடிப்படையில் நேரடியாகக் கணக்கிடப் பட்டது.

மேலும், பெறப்பட்ட சூத்திரத்தின் அடிப்படையில்,

$$V(r) = \frac{1}{(\bar{X})^2}\cdot \frac{N-n}{Nn}(S_y^2 + R^2.S_x^2 - 2\rho S_x S_y)$$

$$= \frac{1}{3^2}\cdot \frac{(5-3)}{5 \times 3}[2{\cdot}5 + 2{\cdot}5 - (2 \times 2{\cdot}0)]$$

$$= \frac{1}{9}\cdot \frac{2}{15}[1{\cdot}0]$$

$$\therefore \quad V(r) = \frac{2}{135} = 0{\cdot}01482$$

இரண்டும் சற்றேறக்குறைய ஒன்றுன மதிப்பாதல் காண்க.

$V(r)$-ன் மாதிரி அளவை மதிப்பீடு :

அதாவது $v(x)$ மதிப்பினை அறிவதாகும். இதற்கு,

$$\hat{V}(r) = v(r)$$

$$= \frac{1}{(\bar{x})^2} \cdot \frac{N-n}{Nn} \cdot \sum_n \frac{(y_i - rx_i)^2}{n-1} \qquad \ldots(A)$$

என்ற விதியினைப் பயன்படுத்தவேண்டும்.

அதாவது,

$$\hat{V}(r) = v(r)$$

$$= \frac{1}{(\bar{x})^2} \cdot \frac{(N-n)}{Nn} (s_x^2 + \hat{R}^2 \cdot s_y^2 - 2\hat{\rho} \cdot \hat{R} \cdot s_x \cdot s_y) \quad \ldots(B)$$

{வெவ்வேறு மாதிரி அளவைகள் பிரதியிடப்படுகின்றன.}

எடுத்துக்காட்டு 3

எடுத்துக்காட்டு (1)-ல் உள்ள 3-வது மாதிரியினை எடுத்துக்கொள்கிறேம். அதில்

y_i	x_i	y_i^2	x_i^2	$y_i x_i$
1	2	1	4	2
2	1	4	1	2
5	5	25	25	25
8	8	30	30	29

(A) அல்லது (B) விதியைப் பயன்படுத்தலாம்.

இங்கு (A) விதியினைப் பயன்படுத்தி.

$$v(r) = \frac{1}{(\bar{x})^2} \cdot \frac{N-n}{Nn} \sum_n \frac{(y_i - r \,.\, x_i)^2}{n-1}, \left[r = \frac{8}{8} = 1 \right]$$

இதில் $\sum_{n} (y_i - rx_i)^2 = \sum y_i^2 + r^2 \sum x_i^2 - 2r \sum x_i y_i$

$$= 30 + 30 - (2 \times 29)$$

$$= 60 - 58 = 2$$

$$\therefore \quad v(r) = \frac{2 \times 3}{(2{\cdot}7)^2 \times 15 \times 2} = \frac{1}{0{\cdot}9 \times 2{\cdot}7 \times 15}$$

$$= \frac{1}{36{\cdot}45}$$

$\therefore \quad v(r) = 0{\cdot}027 \,; \quad \therefore \quad s(r) = 0{\cdot}168.$

மாதிரி அளவு n-ஐத் தீர்மானித்தல் :

$\hat{R}$ ஆகிய r-ன் மாறுபாட்டின் மதிப்பீட்டு அளவை (estimator)

$$\hat{V}(r) = v(r) = \frac{1}{(\bar{x})^2} \cdot \frac{N-n}{Nn} \cdot \sum_{n} \frac{(y_i - rx_i)^2}{n-1}$$

என்று தரப்படுகிறது.

இதில் $\dfrac{\sum_{n} (y_i - rx_i)^2}{(n-1)} = s_d^2$ என்று குறிக்கப்படும்.

ஏதேனும் ஒரு முன்னோடியான மாதிரியினை n_1 என்ற அளவில் எடுத்து, $s_{1d}^2 = \sum_{n_1} \dfrac{(y_i - rx_i)^2}{n_1 - 1}$-யினைக் கணக்கிட்டு, அதனை s_d^2-ன் மதிப்பீடாகக் கொள்க. பின்பு,

$$v(r) = \frac{1}{(\bar{x})^2} \cdot \frac{N-n}{Nn} s_{1d}^2 \text{ என்றாகும்.}$$

z என்ற நம்பகத் தன்மையுடன் (reliability) திட்ட அளவினை 'd' என்று கொண்டு,

$$d^2 = z^2 \cdot v(r)$$

$$= z^2 \cdot \frac{1}{(\bar{x})^2} \cdot \frac{N-n}{Nn} \cdot s_{1d}^2$$

என்று வரையறுக்கிறோம்.

$\frac{N-n}{N} = 1$ என்று கருதி, n-ன் மதிப்பை அறிந்து, அதை n_0 என்று குறிப்பிட,

$d^2 = z^2 \cdot \frac{1}{(\bar{x})^2} \cdot \frac{s_{1d}^2}{n_0}$ என்று பெற்ற மதிப்பிலிருந்து,

$$n_0 = \frac{z^2 \cdot s_{1d}^2}{d^2 \cdot (\bar{x})^2}$$

என்று அறியப்படும்.

பொதுவாக $n_0 = \frac{z^2 \cdot \hat{s}_d^2}{d^2 \cdot (\bar{x})^2}$ (இதில் $\hat{s}_d^2$ என்பது s_d^2-ன் ஏதேனும் முந்தைய மதிப்பீடாகும்.)

எனவே, இறுதியில்

$$\frac{d^2 \cdot (\bar{x})^2}{z^2 \cdot \hat{s}_d^2} = \frac{N-n}{n \cdot N}$$

அதாவது $n_0 = \frac{nN}{N-n}$ என்று தீர்மானிக்கிறோம்.

இதிலிருந்து,

$$n = \frac{n_0}{+\left(\frac{n_0}{N}\right)}$$

என்று நிர்ணயிக்கப்படுகிறது.

இதில் n_0 என்பது, $n_0 = \frac{z^2 \cdot \hat{s}_d^2}{d^2 \cdot (\bar{x})^2}$ என்பதாக பிரதியிடப்படும்.

n_0-உடன் ஒப்பிடுகையில் N அளவு மிகப் பெரிதாக அமைந்திருப்பின், $\frac{n_0}{N}$ என்பதைக் கருதாமல், $n = n_0$ என்றே முடிவு செய்ய இயலும்.

எடுத்துக்காட்டு 4

$r = 0 \cdot 2$ என்பதாக, $n = 10$ என்ற அளவோடு கருதப்பட்ட சம வாய்ப்பு மாதிரியின் அடிப்படையில் கணக்கிடப்படுகிறது.

r-ன் திட்டமானது $(d) \pm 0{\cdot}004$ என்ற வரம்புக்குள் அமைய வேண்டியிருப்பின், $z = 2$ எனக் கொண்டு, மாதிரி அளவினை (n) மதிப்பிடுக. [$\Sigma(y_i - r\, x_i)^2 = 1$ எனக் கொள்க] ; $\bar{x} = 23$.

$n_1 = 10$ என்று கொடுக்கப்பட்டது.

$$\hat{s}_d^2 = \frac{\Sigma\,(y_i - rx_i)^2}{9} = \frac{1}{9}.$$

எனவே வரையறை மூலம்,

$$n_0 = \frac{z^2 \cdot \hat{s}_d^2}{d^2 \cdot (\bar{x})^2}$$

என்ற மாதிரி அளவானது குறிப்பிட்ட திட்பத்தை அளிக்க வல்லதாகும்.

$$\therefore\; n_0 = \frac{4 \times (0{\cdot}11)}{(0{\cdot}004)^2 \times (23)^2} = \frac{0{\cdot}44}{0{\cdot}008} = 55$$

எனவே $n = 10$ என்ற மாதிரி அளவு மிகச் சிறியதாக அமைதல் காணத்தக்கது.

n மதிப்பைத் தீர்மானித்து, அதன்பின் s_d^2-யினைக்கணக்கிட்டு, வேறுபாடு அதிக அளவில் இருப்பின், n_0 மற்றும் n மதிப்புக்கள் திரும்பவும் கணக்கிடப்படவேண்டும். வேறுபாடு குறையும் வரையில் இதனைத் தொடரவேண்டும்.

R-ன் நம்பக இடைவெளி (Confidence Interval) :

n-ன் மதிப்பு 30-க்குமேல் வந்தால், பேரின மாதிரிகளின் (large samples) நியதிப்படி, இயல்நிலைப் பரவலை (normal distribution), r என்ற வாய்ப்பு மாறிக்கு (random variable) கருதி, $\bar{y}$, $\bar{x}$ இவற்றின் மாறுபாட்டுக் கெழுக்கள் (coefficient of variation) 0·1-ஐவிடக் குறைவானதெனக் கொண்டு, R-ன் நம்பக எல்லையானது

$$r \pm z \,.\, s\,(r)$$

என்றபடி வரையறுக்கப்படும்.

கூட்டு மொத்தத்தின் மதிப்பீடு :

$r = \dfrac{\hat{Y}}{X}$ (X கொடுக்கப்பட்டது) என்பதன் மூலம்,

$\hat{Y} = r \cdot X$ என்று மதிப்பிடப்படுகிறது.

இதில் $r = \frac{\frac{\sum y_i}{n}}{\sum x_i}$ என்பதாக ஏற்கெனவே கணிக்கப்பட்டது. X-ன் மொத்த அளவு (total) தெரிந்திருப்பதால் அதனைத் துணை விவரமாகக் (supplementary data) கருதுகிறோம்

சம வாய்ப்பு மாதிரி முறையில்,

$$\hat{Y} = N \cdot \bar{y}$$

என்ற மதிப்பீட்டைப் பெறுகிறோம்.

இவ்விரண்டு மதிப்பீடுகளையும் ஒப்பிடும் பொருட்டு, அவற்றின் மாறுபாட்டு அளவுகளைப் (variances) பின்வருமாறு கருதுவோம் :

விகித மதிப்பீட்டு முறைப்படி,

$$V(\hat{Y}) = V(r \cdot X)$$

$$= N^2 \cdot \frac{N-n}{Nn} \cdot (S_y{}^2 + R^2 \cdot S_x{}^2 - 2\rho\, RS_x\, S_y)$$

என்றும்,

சமவாய்ப்பு மாதிரி முறைப்படி,

$$V_{ran} = V(\hat{Y}_{ran}) = N^2. \frac{N-n}{Nn}. S_y{}^2$$

என்றும் காண்கிறோம்.

இவற்றினிடை உள்ள வேறுபாடு

$$V_{ran} - V(\hat{Y}_R) = N^2. \frac{N-n}{Nn}. (2\rho S_x S_y - R^2.S_x{}^2).$$

{ $\hat{Y}_R$ என்பது விகித மதிப்பீட்டு முறையினைக் குறிக்கும் }.
இதன் மூலம்,

$2\rho RS_xS_y > R^2S_x{}^2$ என்று

அல்லது

$$\rho > \frac{1}{2}. \frac{S_x}{S_y}. R = \frac{1}{2}. \frac{C_x}{C_y}$$

என்று அமையும் நிலையில், $V(\hat{Y}_{ran}) > V(\hat{Y}_R)$ என்று காணப் படுவதாக அறியலாம். அதாவது, V_{ran} மதிப்பினைவிட $V(\hat{Y}_R)$ என்ற மதிப்பீடானது அதிகத் திறன் உடையதென்று அறியப்படு கிறது.

(அல்லது) Y-ன் மாறுபாட்டுக் கெழுவானது (coefficient of variation) X-ன் மாறுபாட்டுக்கெழுவைவிட அதிகமாக இருப்பின், விகித மதிப்பீட்டுமுறை அதிகதிறன் உடையதாய் விளங்கும் (இதற்கு $\rho > 0{\cdot}5$) என்றும் இருக்கவேண்டும். எனவே, மதிப்பீட்டு முறை அதிக நுட்பத்துடன் விளங்க விகித முறை பயன்படுத்தப்படலாம். ஆனால், துணை விவரங்கள் கொடுக்கப்பட்டால் மட்டுமே இது சாத்தியமாகிறது.

பயிர் விளைச்சல் (crops) பற்றிய அளவெடுப்புகளில் (surverys), பயிரிடப்படும் மொத்தப் பரப்பினை முன்கூட்டியே அறிந்து, மாறாததாகக் கருதும் நிலையில், மொத்த விளைச்சலைப் பற்றிய (Y) மதிப்பீட்டினை, விகித முறைப்படி கண்டறிய இயலு கிறது. மக்கட் தொகைக் கணிப்பு(census)பற்றிய ஆய்வுகளிலும், இம் முறையானது அதிகப் பயன் தருகிறது.

பயிற்சி

1. $n = 20$ அளவுள்ள மாதிரித் தெரிவில், கண்டறிந்த விவரங்களாவன :

தங்கு மிடங்கள்	அறைகள் (x_i)	காலியாக உள்ளவை (y_i)	தங்குமிடம்	அறைகள்	காலியாக உள்ளவை (y_i)
1	20	4	11	60	15
2	10	2	12	50	8
3	30	6	13	30	5
4	40	8	14	80	15
5	20	5	15	50	12
6	15	3	16	30	7
7	25	5	17	40	8
8	30	6	18	90	20
9	20	3	19	60	12
10	20	5	20	70	8

காலியாக உள்ள அறைகள் விகிதத்தை மதிப்பிட்டு அறிய வேண்டியுள்ளது. இவ் விவரங்களைப் பயன்படுத்தி,

(i) r-ன் மதிப்பைக் கணக்கிடுக.

(ii) $\frac{N-n}{N} = 1$ என்று கொண்டு, $V(\hat{R})$ன் மதிப்பினைக் காண்க.

(iii) $\pm 0{\cdot}005$ என்ற வரையறையில் திட்பமானது அமைய, $Z = 2$ என்ற நம்பக அளவில், n-ஆனது எவ்வளவாக இருக்கவேண்டும் என்று மதிப்பிடுக.

2.

வரிசை எண்	Y_i	X_i
1	1	3
2	2	1
3	3	2
4	4	5
5	5	4

மேற்கண்ட பட்டியலை உபயோகித்து, $(N = 5)$,

(i) R-ன் மதிப்பினைக் காண்க

(ii) $n = 3$ என்ற அளவில் உள்ள எல்லாவழி மாதிரிகளையும், திரும்ப வைக்காத முறையில் கருதி, r-ன் மதிப்புகளை அறிக.

(iii) பிறழ்ச்சியின் மதிப்பு என்ன ?

(iv) $V(r)$-ன் மதிப்பினையும் கணக்கிடுக.

(ஆ) படுகை மாதிரிமுறையும், விகித மதிப்பீடும் :

முழுத் தொகுதியானது பல படுகைகளாகப் பிரிக்கப்பட்டு, R என்ற தொகுதி விகிதமானது மதிப்பிடப்படுகின்றது.

$$\hat{R} = \frac{\hat{Y}}{\hat{X}} = \frac{y}{x} = \frac{\bar{y}}{\bar{x}}$$

என்பதாக மதிப்பீடு காணப்படும். அதற்கு $\hat{Y}$, $\hat{X}$ இவற்றைப்

பெறவேண்டும். இதற்குப் படுகை மாதிரிமுறை பயன்படுத்தப் படுகிறது.

இதற்கான இரு வழிமுறைகள் :

(i) சேர்வு அல்லது இணைத்த முறை (Combined)

(ii) தனித்த முறை (Separate estimate).

(i) **சேர்வுமுறை : (இணைத்த முறை) :**

R-ன் மதிப்பீடு :

இதில், Y, X இவற்றை மதிப்பீடு செய்து, r-ன் மதிப்பைப் பெறுகிறோம். படுகை முறைப்படி (stratified),

$$\hat{Y}_{st} = \sum_{h=1}^{L} \frac{N_h}{n_h} \sum_{1}^{n_h} y_{hi} = \Sigma N_h \bar{y}_h$$

$$\hat{X}_{st} = \sum_{1}^{L} \frac{N_h}{n_h} \sum_{1}^{n_h} x_{hi} = \Sigma N_h . \bar{x}_h$$

என்று பெறுவதால்,

$$\hat{R} = r_c = \frac{\hat{Y}_{st}}{\hat{X}_{st}} = \frac{\sum_{L} N_h/n_h \sum_{n_h} y_{hi}}{\sum_{L} N_h/n_h \sum_{n_h} x_{hi}} = \frac{\Sigma N_h \bar{y}_h}{\Sigma N_h \bar{x}_h}$$

என்று மதிப்பிடப்படும். (r_c என்பது, இணைப்பு முறையில் r-ன் மதிப்பினைக் குறிப்பது.)

மேலும், இம் மதிப்பீட்டின் மாதிரிமுறை சார்ந்த மாறுபாட்டு அளவு

$$V(\hat{R}) = V(r_c)$$

$$= \frac{1}{N^2} . \frac{1}{(\bar{X})^2} . \sum_{L} N_h^2 . \frac{N_h - n_h}{N_h} . \frac{1}{n_h} .$$

$$(S_{yh}^2 - 2R\rho \cdot S_{xh} \cdot S_{yh} + R^2 \cdot S_{xh}^2)$$

$$= \frac{1}{(X)^2} \sum_{L} N_h^2 . \frac{N_h - n_h}{N_h} . \frac{1}{n_h} . S_{ch}^2$$

என்று பெறப்படும்.

இதில், $S_{ch}^2 = (S_{yh}^2 - 2R \cdot \rho \cdot S_{xh} \cdot S_{yh} + R^2 \cdot S_{xh}^2)$;

$$S_{yh}^2 = \frac{1}{N_h - 1} \Sigma (Y_{hi} - \bar{Y}_h)^2 \text{ (படுகைமுறை) };$$

$$\rho_h = \frac{S_{xyh}}{S_{xh} \cdot S_{yh}}$$

என்று அமைகின்றன.

[ρ_h என்ற ஒவ்வொரு படுகையின் ஒட்டுறவுக் கெழுவும் தேவைப்படுவதை அறிக.]

ρ_h அதிகமாக இருக்கும் நிலையில் $(R > 0)$, $V(r_c)$ யானது மிகவும் குறைகின்றது.

மாதிரி அளவைகளான s_{yh}, s_{xh}, $\hat{R}_1$ $\hat{\rho}_h$ இவற்றைப் பயன் படுத்தி,

$$\hat{V}(r_c) = \frac{1}{N^2} \cdot \frac{1}{(\hat{X})^2} \sum_{L} N_h^2 . \frac{N_h - n_h}{N_h} . \frac{1}{n_h}$$

$$(s_{yh}^2 - 2r_c \cdot \rho_h \cdot s_{xh} \cdot s_{yh} + r_c^2 \cdot s_{xh}^2)$$

என்று மதிப்பிட்டு அறிகிறோம்.

எடுத்துக்காட்டு 1 : முழுமைத் தொகுதி

	N_h	n_h	Y_{hi}	X_{hi}
படுகை 1	5	2	1, 2, 3, 3, 4	6, 7, 7, 8, 10
படுகை 2	4	2	3, 4, 4, 5	14, 14, 18, 19
படுகை 3	4	2	5, 6, 8, 8	30, 33, 40, 48

(i) $R = \frac{Y}{X} = \frac{56}{254} = 0 \cdot 22$

(ii) ஒவ்வொரு படுகையிலிருந்தும் $n_h = 2$ என்ற அளவில் மாதிரிகளை (சமவாய்ப்புடன்) எடுக்கலாம். உதாரணமாக,

	[illegible]	y_{hi}	y_h	$\bar{y}_h$	x_{hi}	x_h	$\bar{x}_h$
படுகை 1	2	1, 2	3	1·5	6, 7	13	6·5
,, 2	2	3, 4	7	3·5	14, 14	28	28
,, 3	2	5, 6	11	5.5	30, 33	63	31·5
	6		21			104	

என்று கொள்வோம்.

$$\hat{Y} = \Sigma N_h \cdot \bar{y}_h = (5 \times 1{\cdot}5) + (4 \times 3{\cdot}5) + (4 \times 5{\cdot}5)$$

$$= 7{\cdot}5 + 14{\cdot}0 + 22{\cdot}0 = 43{\cdot}5$$

$$\hat{X} = \Sigma N_h \bar{x}_h = 270{\cdot}5$$

$$\therefore \quad \hat{R} = r_c = \frac{43{\cdot}5}{270{\cdot}5} = 0{\cdot}16$$

(iii) $V(\hat{R}) = V(r_c)$ மதிப்பீடு :

$(n_h = 2)$.

பட்டியலானது :

படுகை	S_{yh}^2	S_{xh}^2	$R^2 \cdot S_{xh}^2$	$2R \cdot \rho_h \cdot S_{xh} \cdot S_{yh}$	S_{ch}^2
1	1·3	2·3	0·1113	0·5456	0·87
2	0·67	6·9	0·3340	0·55	0·45
3	0·73	64·25	3·1100	3·6036	1·24

$$\bar{Y}_1 = 2{\cdot}6 \; ; \quad \bar{Y}_2 = 4 \; ; \quad \bar{Y}_3 = 6{\cdot}75$$

$$_1 = 7{\cdot}6 \; ; \quad \bar{X}_2 = 16{\cdot}25 \; ; \quad \bar{X}_3 = 37{\cdot}75$$

$$R = 0{\cdot}22 \qquad \therefore \; R^2 = 0{\cdot}0484$$

$\rho_1 = \dfrac{S_{xy1}}{S_{x1} \cdot S_{y1}}$ என்றவாறு பெறப்படும்.

இதேபோல் ρ_2, ρ_3 ஆகியவற்றையும் மதிப்பிடலாம்.

$S_{xy1} = 1{\cdot}24$; $S_{xy2} = 1{\cdot}25$; $S_{xy3} = 8{\cdot}19$

S_{ch}^2 மதிப்பு மூன்று படுகைகட்கும் அறியப்படுகிறது.

$$\therefore \quad S_{ch}^2 = (S_{yh}^2 + R^2 \cdot S_{xh}^2 - 2R \cdot \rho_h S_{xh} S_{yh})(h=1, 2, 3)$$

இவற்றினின்று,

$$V(r_c) = \frac{1}{(\bar{X})^2} \sum_L N_h^2 \cdot \frac{N_h - n_h}{N_h} \cdot \frac{1}{n_h} \cdot S_{ch}^2$$

$$= \frac{1}{(254)^2} \left[\frac{25 \times 3}{5} \times \frac{(0{\cdot}87)}{2} + \frac{16 \times 2}{4} (0{\cdot}45 + 1{\cdot}24) \right]$$

$$\therefore \quad V(r_c) = \frac{1}{64516} [20{\cdot}045] = 0{\cdot}00033$$

என்று மதிப்பிடப்படுகிறது.

∴ நடைமுறைக்கு இம்முறை எளிதாக அமைவதில்லை.

பயிற்சி : கருதப்பட்ட மாதிரியினைப் பயன்படுத்தி $\hat{V}(r_c)$-யினை மதிப்பிடுக.

Y என்ற மொத்தத்தினை (total),

$\hat{Y}_c = r_c \cdot X$ என்று மதிப்பிடலாம். இதற்கு X என்ற மதிப்பு ஏற்கெனவே தெரிந்ததாக இருத்தல் அவசியம்.

$$\text{மேலும்} \quad V(\hat{Y}_c) = \Sigma N_h^2 \cdot \frac{N_h - n_h}{N_h} \cdot \frac{1}{n_h} \cdot S_{ch}^2$$

என்று மதிப்பிடப்படும்.

இனி இரண்டாவது முறையினைக் கருதுகிறோம்.

(ii) தனித்த மதிப்பீட்டு முறை :

X-ன் மதிப்பை அறிந்திருப்பதாகக் கொண்டு $r = \frac{\hat{Y}}{X}$ என்ற விகிதத்தில், தொகுதி உறுப்பான $\hat{Y}$-யினை, விகிதமதிப்பீட்டு

முறையினை ஒவ்வொரு படுகையிலும் கருதி, மதிப்பீடு செய்கிறோம்.

h என்ற படுகையில்

$$r_h = \frac{y_h}{x_h}$$

என்று கொண்டால்,

$$\hat{Y}_h = r_h \cdot X_h$$

$$\therefore \quad \hat{Y}_h = \frac{y_h}{x_h} \cdot X_h$$

என்பது, h-ஆவது படுகையின் மொத்தத்தின் மதிப்பீடாகப் பெறப்படும்.

எனவே $\hat{Y}_s = \Sigma \hat{Y}_h = \Sigma r_h \,.\, X_h$
என்று மொத்த அளவான Y-ஆனது மதிப்பீடு செய்யப்படுகிறது. ($\hat{Y}_s$ என்பது தனித்த முறையினைக் குறிக்கிறது.) எனவே, X-ன் மதிப்பு ஏற்கெனவே கொடுக்கப்பட,

$$r_s = \Sigma \frac{X_h}{X} r_h$$

என்று பெறப்படுகிறது.

மேலும்

$$V(r_s) = \frac{1}{N^2} \cdot \frac{1}{(\bar{X})^2} \Sigma N_h^2 \cdot \frac{N_h - n_h}{N_h} \cdot \frac{1}{n_h} \cdot S_{sh}^2$$

என்று பெறப்படும்.

இதில் $S_{sh}^2 = (S_{yh}^2 - 2R_h \cdot \rho_h \cdot S_{xh} \cdot S_{yh} + R_h^2 \cdot S_{xh}^2)$ என்பதைக் காண்க.

எடுத்துக்காட்டு 2 :

எடுத்துக்காட்டு (1)-ல் இருந்து, $n_h = 2$ என்று ஒவ்வொரு படுகையிலிருந்தும் மாதிரிகள் எடுக்க.

y_{hi}	y_h	$\bar{y}_h$	x_{hi}	x_h	$\bar{x}_h$	$r_h = \frac{y_h}{x_h}$
1, 3	4	2	6, 8	14	7	$\frac{4}{14} = 0{\cdot}29$
3, 5	8	4	14, 19	33	16·5	$\frac{8}{33} = 0{\cdot}24$
8, 8	16	8	40, 48	88	44	$\frac{16}{88} = 0{\cdot}18$

r_h மதிப்புகளினின்று $\hat{Y}_s$ மற்றும் r_s காணுதல் :

X_h	$\hat{Y}_h = r_h \cdot X_h$
38	11·02
65	15·6
151	27·2

எனவே, $\hat{Y}_s = \Sigma\, r_h \cdot X_h = 53{\cdot}82.$

$$\therefore \quad r_s = \frac{53{\cdot}82}{254} = 0{\cdot}22$$

இனி $V(r_s)$-ஐ மதிப்பிடல் :

இதற்கு S_{sh}^2-மதிப்பை ஒவ்வொரு படுகைக்கும் கணக்கிட வேண்டும். இதற்கான பட்டியல் (முழுமைத் தொகுதியைக் கருதி) :

படுகை	S_{yh}^2	S_{xh}^2	$R_h^2 \cdot S_{xh}^2$	$2R_h \cdot \rho_h \cdot S_{xh} \cdot S_{yh}$	S_{sh}^2
1	1·3	2·3	0·069	0·744	0·625
2	0·67	6·9	0·276	0·5	0·446
3	1·73	64·25	1·9275	2·9484	1·012

$$R_1 = \frac{Y_1}{X_1} = \frac{13}{38} = 0{\cdot}3 \qquad \therefore \quad R_1^2 = 0{\cdot}09$$

$$R_2 = \frac{16}{65} = 0{\cdot}2\,; \qquad \therefore \quad R_2^2 = 0{\cdot}04$$

$$R_3 = \frac{27}{151} = 0{\cdot}18 \qquad \therefore \quad R_3^2 = 0{\cdot}03$$

$$\therefore \quad V(r_s) = \frac{1}{(254)^2}\left[\frac{75 \times (0{\cdot}625)}{10} + \frac{16 \times 2}{4}(1{\cdot}42)\right]$$

$$\therefore \quad V(r_s) = \frac{16{\cdot}05}{64516} = 0\ 0003.$$

Y-ன் மதிப்பீடு :

$$\hat{Y}_s = \Sigma \hat{Y}_h = \Sigma \frac{y_h}{x_h} \cdot X_h$$

என்பதும்,

$$V(\hat{Y}_s) = V(r_s \cdot X) = X^2 \cdot V(r_s)$$

$$= \Sigma N_h^2 \cdot \frac{(N_h - n_h)}{N_h} \cdot \frac{1}{n_h} \cdot S_{sh}^2$$

என்பதும், தனித்த முறைப்படி பெறப்பட்ட Y என்ற மொத்தம் பற்றிய அளவீடுகள் ஆகும்.

எடுத்துக்காட்டு 3 :

எடுத்துக்காட்டு (2)-ல்

$$\hat{Y}_s = \Sigma r_h \cdot X_h = \Sigma \hat{Y}_h = 54$$

$$\therefore \quad V(\hat{Y}_s) = X^2\, V(r_s)$$

$V(r_s) = 0{\cdot}0003$ என்று அறியப்பட்டது.

எனவே

$$V(\hat{Y}_s) = (254)^2 \times 0{\cdot}0003$$

$$= (254)^2 \left[\frac{1}{(254)^2} \times (16{\cdot}05)\right]$$

$$\therefore \quad V(\hat{Y}_s) = 16{\cdot}05$$

என்றும்,

$$\sigma(\hat{Y}_s) = \sqrt{16{\cdot}05} \doteqdot 4{\cdot}0$$

என்றும் மதிப்பிடப்படுகின்றன.

பயிற்சி

1. பின்வரும் முழுமைத் தொகுதியினைக் கருதி, கேள்விகட்கு தக்க பதில்களைத் தரவும் :

(விற்பனை : ரூ. 1000-ல்)

கடைகள்	N_h	Y_{hi} (மே மாத விற்பனை)	X_{hi} ஏப்ரல் விற்பனை
சிறியன	10	2,4,1,3,1,5,4,4,5,3	3,4,2,4,2,5,4,5,6,5
நடுத்தரம்	8	5,6,5,7,8,8,9,11	6,8,6,9,11,8,10,13
பெரியவை	5	11,13,15,12,18	13,12,18,14,24.

(i) R என்ற முழுமைத் தொகுதி விகிதத்தைக் காண்க. {மே-ஏப்ரல் விற்பனை விகிதம்}.

(ii) $n_h = 3$ என்ற அளவில் ஒரு மாதிரியினை ஒவ்வொரு படுகையினின்றும் தெரிவு செய்து, R-ன் மதிப்பீடு காண்க.

(iii) மேற்கண்ட மாதிரி அடிப்படையில், $\hat{V}(r_c)$ மற்றும் $s(r_c)$ மதிப்புகளைக் கணக்கிடுக.

(iv) Y மற்றும் X மதிப்புகளைக் கணக்கிடுக. மேலும் $\hat{Y}_c$-யின் மதிப்பீடு என்ன ?

(v) $\hat{V}(\hat{Y}_c)$ மற்றும் $s(\hat{Y}_c)$ இவற்றைக் கணக்கீடு செய்க.

2. பயிற்சி (1)-ல் கொடுக்கப்பட்ட அட்டவணையினைப் பயன்படுத்தி, பின்வருவனவற்றைக் கணக்கிடுக :

(i) ஒவ்வொரு படுகையினின்றும் $n_h = 3$ என்றபடி மாதிரிகள் எடுத்து r_s-ன் மதிப்பினை அறிக.

(ii) $V(r_s)$ மற்றும் $\sigma(r_s)$ இவற்றைக் காண்க.

(iii) $\hat{V}(r_s)$ மற்றும் $s(r_s)$ இவற்றினை அறிக.

(iv) $\hat{Y}_s$-ன் மதிப்பினைப் பெற்று, $\hat{V}(\hat{Y}_s)$-ஐ தீர்மானிக்க.

மாதிரி பங்கீட்டு முறை
(Allocation of Sample)

n என்ற மாதிரி அளவு தீர்மானிக்கப்படுவதாகக் கொண்டு, அதனை (n-ஐ) வெவ்வேறு படுகைகளில் எவ்வாறு பங்கீடு செய்வது ?

உத்தமப் பங்கீட்டு முறையின்படி, கொடுக்கப்படும் செலவு நிபந்தனைக்குட்பட்டு, மாறுபாட்டினைச் சிறும அளவினதாக்கத் தக்க அளவில் பங்கீட்டு முறை அமைதல் வேண்டும்.

அதாவது, $C = C_0 + \Sigma\ n_h C_h$ என்ற செலவு நிபந்தனைக் கேற்ப,

$$V(\bar{y}_{st}) = \frac{1}{N_2} \Sigma N_h^2 . \frac{N_h - h_h}{N_h} . \frac{S_h^2}{n_h}$$

ஆனது சிறும அளவாக்கப்பட்ட (minimised) வேண்டும்.

மேலும்

$$n_h = \frac{N_h S_h / \sqrt{C_h}}{\Sigma N_h S_h / \sqrt{C_h}} . n \qquad \ldots\ldots (1)$$

என்று பெறப்பட்டது.

C_h என்பது மாறுதிருப்பின்,

$$n_h = \frac{N_h S_h}{\Sigma N_h S_h} . n \qquad \ldots\ldots (2)$$

என்று 'நேமன் பங்கீட்டு முறை' பெறப்பட்டது.

(2)-ஆவது முறையின் அடிப்படையில் முடிவினை எளிதில் பெறலாம்.

முதலில் இணைத்த விகித மதிப்பீட்டில் கருதுவோம். இதன்படி,

$$V(r_c) = \frac{1}{N^2} \frac{1}{(\bar{X})^2} \Sigma\ N_h^2 . \frac{N_h - h_h}{N_h} . \frac{1}{n_h} . S_{ch}^2$$

என்றும்

$$V(\hat{Y}_c) = \Sigma N_h^2 . \frac{N_h - n_h}{N_h} . S_{ch}^2$$

என்றும் பெறப்பட்டன.

எனவே, (2)-ல், $S_h = \frac{S_{ch}}{\bar{X}}$

என்று பிரதியிட,

$$n_h = \frac{N_h (S_{ch}/\bar{X})}{\sum_h N_h (S_{ch}/\bar{X})} \cdot n = \frac{N_h S_{ch}}{\sum_h N_h S_{ch}} \cdot n$$

என்று மதிப்பினைக் காண்கிறோம்.

மேலும், $\hat{Y}_c$ யினைக் கருதுகையிலும்,இதே மதிப்பு பெறப்படுகிறது.

இதன்மூலம், சிறு அளவினதாகவும், S_{ch}^2 மூலம் தரப்படும் மாறுபாடானது சிறிதாகவும் அமைந்திருக்கும் படுகைகளிலிருந்து, சிறிய அளவில் மாதிரிகள் தெரிவு செய்யப்படுதல் வேண்டும் என்ற கருத்து புலனாகிறது. ஆனால், S_{yh}, S_{xh} ஆகியவை சிறிதாகவும், ρ_h-ன் மதிப்பு ஒன்றுக்கு (1) அண்மையிலும் இருக்கும் நிலையில்,S_{ch}^2 மதிப்பும் சிறியதாக இருக்கும். அதாவது, S_{yh}^2, S_{xh}^2, ρ_h இவற்றை மதிப்பீடு செய்யும் பொருட்டு எடுக்கப் படும் மாதிரியானது சிறிய அளவினதாதல் வேண்டும்.

அடுத்து, தனித்த விகித மதிப்பீட்டில்,

$$V(r_s) = \frac{1}{N^2} \frac{1}{(\bar{X})^2} \Sigma N_h^2 . \frac{N_h - n_h}{N_h} . \frac{1}{n_h} . S_{sh}^2 .$$

மேலும், n_h மதிப்பு, r_s $\hat{Y}_s$ இவற்றிற்கு ஒன்றாகவே அமைந்து,

$$n_h = \frac{N_h \cdot S_{sh}}{\Sigma N_h \cdot S_{sh}} . n$$

என்றாகும்.

இதில் ஏற்பட்ட மாற்றம், S_{ch} க்கு பதிலாக, S_{sh} அமைவதே முன் கூறப்பட்ட கருத்துக்கள் இதற்கும் பொருந்துவனவாகும்.

மாதிரி அளவு (n) :

படுகை மாதிரியில் கருதப்பட்ட அதே வழிமுறையில்,

$n_h = \frac{N_h}{N} . n$ எனக் கொண்டு,

$$V\left(\frac{\hat{Y}_c}{N}\right) = \frac{1}{N^2} . \Sigma N_h^2 . \frac{N_h - n_h}{N_h} . \frac{1}{n_h} . S_{ch}^2$$

என்றும் அறிந்து,

$$n = \frac{N \Sigma N_h \cdot S_{ch}^2}{N^2 D^2 + \Sigma N_h \cdot S_{ch}^2}$$

என்று பெறுகிறோம்.

$$\left(D^2 = \frac{d^2}{z^2}\right).$$

இணைத்த முறை மற்றும் தனித்த முறை பற்றிய ஒப்புமை :

இருமுறைகளில், குறைந்த அளவு பிறழ்ச்சியும், சிறிய மாறுபாட்டு அளவும், மற்றும் சில பயன்களும் அமையும் விகித முறை சிறந்த தெனக கருதப்படும்.

பிறழ்ச்சி :

தனித்த விகித மதிப்பீட்டில் அதிக அளவில் காணப்படும். குறிப்பாக, ஒவ்வொரு படுகையிலும் n_h-ஆனது சிறிதாக இருப்பின், இது மிகவும் பொருந்தும். மொத்தமாக கருதப்படும் மாதிரி அளவானது சிறிதாக இருந்தபோதிலும், ஒவ்வொரு படுகையிலும் உள்ள விகிதங்களின் பிறழ்ச்சிகள் படிப்படியாக அதிகரித்து, இறுதியில் பெறப்படும் விகிதத்தில் உள்ள பிறழ்ச்சியானது அதிகரிக்க நேரிடுகிறது. எனவே, n_h சிறிதாயும், $n = \Sigma n_h$ என்பது பெரியதாகவும் இருக்கும் நிலையில், இணைப்பு விகித மதிப்பீடு குறைந்த அளவு பிறழ்ச்சி உடையதாகக் காணப்படும்.

n_h சிறிதாகவும், ஒவ்வொரு படுகையின் விகிதங்களின் பிறழ்ச்சிகளும் குறைந்ததாகவும் இருப்பின், தனித்த முறை பயன்படுத்தப் பெறலாம். இதற்கு, பின்னடைவுக் கோடானது, ஒவ்வொரு படுகைக்கும், (0, 0) மூலமாக செல்லுமாறு அமைதல் வேண்டும்.

மாறுபாடு :

ஒவ்வொரு படுகைக்கும் n_h மதிப்பு பெரிதாக இருந்தால், $V(r_s)$ என்பது அர்த்தமுள்ளதாக அமையும். எனவே, n_h எல்லாம் சிறியதாகவே இருப்பின், இணைப்பு விகித முறை பயன்படுத்தப்படலாம். மேலும், படுகைகளிடையே R_h ஆனது அதிக அளவில் மாறுபடுகையில், $V(r_s)$ உடன் ஒப்பிடுகையில், $V(r_c)$ யானது குறைந்ததாக இருக்கும்.

இவ்விரு மாறுபாடுகளிடையே உள்ள இடைவெளி

$$V(r_c) - V(r_s) = \frac{1}{(X)^2} \Sigma N_h^2 \cdot \frac{N_h - n_h}{N_h} \cdot \frac{1}{n_h}$$

$$[(R^2 - R_h)^2 S_{xh}^2 - 2(R - R_h) \rho_h S_{xh} \cdot S_{yh})]$$

$$= \frac{1}{(X)^2} \Sigma N_h^2 \cdot \frac{N_h - n_h}{N_h} \cdot \frac{1}{n_h}$$

$$[(R - R_h)^2 S_{xh}^2 + 2(R - R_h) \cdot (2R_h \cdot S_{xh}^2 - 2\rho_h \cdot S_{xh} \cdot S_{yh})]$$

$R \cdot S_x = \rho S_y$ என்னும் நிலையில், பின்னடைவுக் கோடானது (regression line) (0, 0) வழியாகச் செல்ல, பிறழ்ச்சி பூச்சியமாகிறது. இது ஒவ்வொரு படுகைக்கும் அமைகையில்,

$2 R_h \cdot S_{xh}^2 - 2 \rho_h \cdot S_{xh} \cdot S_{yh} = 0$ ஆகிறது

எனவே, $[V(r_c) - V(r_s)] \propto (R - R_h)^2$

அதாவது $V(r_c) - V(r_s)$ என்ற வேறுபாடு, $(R - R_h)^2$ மதிப்புடன் வீதத் தொடர்பு (proportional) உள்ளதாக அமைகிறது.

அதன்படி, படுகைக்குப் படுகை R_h ஆனது வேறு படுகையில், $(R - R_h)^2$ அதிகரிக்க, $V(r_c)$ யானது $V(r_s)$-ஐ விட அதிகமாக இருப்பதைக் காணலாம். அதாவது, அத்தகைய சூழ்நிலையில், தனித்த விகித மதிப்பீட்டினைப் பயன்படுத்துவது திறன் வாய்ந்ததாகும்.

மேலும் $V(r_s)$ கணக்கிடுவதற்கு, ஒவ்வொரு படுகையிலும் r_h-ன் மதிப்பு காணப்படவேண்டும். இந்த r_h மதிப்பு பல கணிப்பு ஆய்வுகளில் விரும்பப்படலாம்.

இவ்வாறு, இருவகையான முறைகளும், படுகை மாதிரி முறை சார்ந்த நிலையில் ஒப்பிடப்படுகின்றன.

(இ) ஒழுங்கு மாதிரி முறையும், விகித மதிப்பீடும் : (Systematic Sampling & Ratio Estimate)

ஒழுங்கு முறையில் மாதிரிக் கூறு தெரிவு செய்யப்படும் நிலையில் பிறழ்ச்சி மற்றும் மாறுபட்டு மதிப்புக்களை விகித மதிப்பீட்டு அளவைகட்கு, தக்க பிரதியிடுதல் மூலமாகப்

பெறலாம். மதிப்பீட்டு அளவைகளான $\bar{x}$, $\bar{y}$ இவற்றின் மாறுபாட்டு மற்றும் உடன் மாறுபாட்டு (covariance) மதிப்புகளைக் கருத வேண்டும்.

ஒழுங்கு மாதிரி முறையில்,

$$V(\bar{y}) = \frac{1}{k}\sum_{i=1}^{k}(\bar{y}_i - \bar{Y})^2;$$

$$V(\bar{x}) = \frac{1}{k}\Sigma(\bar{x}_i - \bar{X})^2$$

மேலும் $\text{Cov}(\bar{y}, \bar{x}) = \frac{1}{k}\Sigma\ (y_i - \bar{y})\ (x_i - \bar{x})$

என்று பெறுகிறோம்.

$\bar{x}_i$, $\bar{y}_i$ ஆகியவை i-ஆவது மாதிரியில் பெறக்கூடிய சராசரி மதிப்புகள். k-என்பது மாதிரி எடுத்தலில் கருதப்பட்ட இடைவெளியாகும் ($nk = N$). இவற்றின் மூலம்,

பிறழ்ச்சி $= B(\hat{R}) = B(r)$

$$= \frac{1}{(\bar{X})^2}\cdot\frac{1}{k}\sum_{i=1}^{k}\bar{x}_i\ (R\cdot\bar{x}_i - \bar{y}_i)$$

என்றும்,

மாறுபாட்டு மதிப்பீடு

$$V(\hat{R}) = V(r)$$

$$\therefore\ V(r) = \frac{1}{(\bar{X})^2}\sum_{i=1}^{k}(\bar{y}_i - R\ .\ \bar{x}_i)^2$$

என்றும் வருவதைக் காணலாம். பொதுவாகக் கொடுக்கப்பட்ட நியதிகளில் தக்கபடி பிரதியிட்டு, இவை பெறப்பட்டன.

V $(\hat{R})$-ன் தொடரிலிருந்து (expression) $\frac{\bar{y}_i}{\bar{x}_i}$ என்ற விகிதங்களின் இடைப்பட்ட மாறுபாடு (between variation) சிறும அளவினதாக அமையும் விதத்தில் உறுப்புகளை (units) அமைப்பதன் மூலமாக, விகித மதிப்பீடானது அதிகத் திறனுடையதாகக் காணலாம்.

ஒழுங்கு மாதிரிமுறையில் கருதப்படும் பிரிவு உள்ளீட்டு ஒட்டுறவுக் கெழு (intra-class correlation) மதிப்பான ρ_{wy}-யினைப் பயன்படுத்தி,

$$V(\bar{y}_{sy}) = \frac{S_y^2}{n} \cdot \left(\frac{N-1}{N}\right) (1 + \overline{n-1} \cdot \rho_{wy})$$

என்று எழுதுகிறோம். இதேபோல் $V(\bar{x}_{sy})$ மற்றும்

$\text{Cov}_{sy}(\bar{x}, \bar{y})$ இவற்றையும் பிரதியிட,

$$V(\hat{R}_{sy}) = \frac{1}{(\bar{X})^2} \cdot \frac{N-1}{N \cdot n} [[S_y^2 (1 + \overline{n-1}\, \rho_{wy})$$

$$+ R^2 \cdot S_x^2 (1 + \overline{n-1} \cdot \rho_{wx})$$

$$- 2R \cdot \rho \cdot S_x \cdot S_y \cdot \sqrt{(1 + \overline{n-1} \cdot \rho_{wx}) (1 + \overline{n-1} \cdot \rho_{wy})}]$$

என்றாகிறது. (ρ என்பது x, y இவற்றினிடையே உள்ள உடன் தொடர்புக் கெழுவாகும்.)

குறிப்பாக, $\rho_{wx} = \rho_{wy} =$ | எனும் நிலையில்,

$$V(\hat{R}_{sy}) = \frac{1}{(\bar{X})^2} \cdot \frac{N-1}{Nn} \cdot [(S_y^2 - 2R\, \rho\, S_x S_y + R^2 S_x^2)$$

$$\times (1 + \overline{n-1} \cdot \rho_w)]$$

அல்லது $V(\hat{R}_{sy}) = \frac{R^2}{n} \cdot \{ C_y^2 - 2\rho\, C_x C_y + C_x^2 \}$

$$\cdot \{1 + (n-1) \cdot \rho_w\}$$

என்ற விதத்தில், மாறுபாட்டுக் கெழுக்களான C_x, C_y இவற்றைப் பயன்படுத்தி, வேறு விதத்திலும் எழுதலாம்.

இதனைச் சம வாய்ப்பு முறையுடன் ஒப்பிட,

$$V(\hat{R}_{sy}) = V(\hat{R}_{srs}) \{1 + (n-1) \cdot \rho_w\}$$

என்று பெறப்படும்.

இதிலிருந்து, விகித மதிப்பீட்டிற்கு, ஒழுங்கு மாதிரிமுறை யானது, ρ மதிப்பு எதிர்க் கணியமாக (—) வரும் விதத்திற் கேற்பக் கூறுகளை அமைத்துக்கொள்வதன்மூலம் திறன் அதிகம் உள்ளதாக இருக்கும்.

(ஈ) திரள் மாதிரித் தேர்வும், விகித மதிப்பீடும் :

திரள் மாதிரிமுறைத் தேர்வில் பெறப்பட்ட முடிவுகளைப் பயன்படுத்தி, விகித மதிப்பீட்டில் அமைகின்ற நிலையினை அறிகிறோம்.

எளிய திரள் முறையில்,

$$\hat{Y}_{cl} = \frac{M}{m} \sum_{1}^{m} \frac{N_i}{n_i} \sum_{1}^{n_i} y_{ij}$$

என்றும்,

$$V(\hat{Y}_{cl}) = M^2 \cdot \frac{M-m}{M} \cdot \frac{S_b^2}{m} + \frac{M}{m} \cdot \sum_{M} N_i^2 \cdot \frac{N_i - n_i}{N_i} \cdot \frac{S_i^2}{n_i}$$

என்றும் கண்டறியப்பட்டன.

$$\left[S_b^2 = \frac{1}{M-1} \sum_{M} (Y_i - \bar{Y})^2 \; ; \quad S_i^2 = \frac{1}{N_i - 1} \sum_{N_i} (y_{ij} - \bar{Y}_i)^2 \right]$$

மேலும், $\hat{\bar{Y}}_{cl} = \frac{\hat{Y}_{cl}}{N}$; $V(\hat{\bar{Y}}_{cl}) = \frac{1}{N^2} \cdot V(\hat{Y}_{cl})$ என்றும் அறியப்பட்டன.

இனி விகித மதிப்பீட்டு நிலையினைக் காண்போம் :

R என்ற விகிதத்தின் மதிப்பீடு

$$r_{cl} = \frac{\hat{Y}_{cl}}{\hat{X}_{cl}} \qquad \text{...(1)}$$

என்று பெறப்படும்.

$$\text{மேலும் } V(r_{cl}) = \frac{1}{\bar{X}^2}\left[M^2 \cdot \frac{M-m}{M} \cdot \frac{S_{1b}^2}{m} + \frac{M}{m} \cdot \sum_{M} N_i^2 \cdot \frac{N_i - n_i}{N_i} \cdot \frac{S_{2i}^2}{n_i} \right] \quad \ldots(2)$$

என்று காணப்படும். இதில்

$$S_{1b}^2 = S_{yb}^2 + R^2 S_{xb}^2 - 2R \cdot \rho_b \cdot S_{xb} \cdot S_{yb} ;$$

$$S_{2i}^2 = S_{yi}^2 + R^2 S_{xi}^2 - 2R \cdot \rho_i \cdot S_{xi} \cdot S_{yi}$$

என்பன ஆகும். இதனுள்,

$$S_{yb}^2 = \frac{1}{M-1} \sum_{M} (Y_i - \bar{Y})^2$$

$$S_{yi}^2 = \frac{1}{N_i - 1} \sum_{N_i} (y_{ij} - \bar{Y})^2$$

எனவும்,

$$\rho_b = \frac{1}{M-1} \sum_{m} \frac{(Y_i - \bar{Y})(X_i - \bar{X})}{S_{xb} \cdot S_{yb}}$$

$$\rho_i = \frac{1}{N_i - 1} \sum_{N_i} \frac{(y_{ij} - \bar{Y}_i)(x_{ij} - \bar{X}_i)}{S_{xi} S_{yi}}$$

எனவும் இருப்பதைக் காணலாம்.

$M = m$ என்னும் நிலையில், (2) ஆனது படுகை மாதிரியில் உள்ள இணைத்த விகித மதிப்பீட்டு நிலைக்குச் சமமாகிறது.

தக்கபடி s_{1b}^2, s_{2i}^2, s_{yb}^2, s_{yi}^2, $\hat{\rho}_b$, $\hat{\rho}_i$ போன்ற மாதிரி மதிப்பீடுகளைப் பிரதியிட்டு, {(2)-ல்}, $V(r_{cl})$-ன் மாதிரி மதிப்பீட்டினை அறியலாம்.

$\frac{m}{M}$ என்பது பூச்சியமாகக் கருதத்தக்க அளவிற்கு M ஆனது m-ஐவிட அதிகமாக இருப்பின்,

$$\hat{V}(r_{cl}) = \frac{1}{\bar{X}^2} \cdot M^2 \cdot \frac{s_{1b}^2}{m}$$

என்ற எளிய அளவில் சுருங்குகிறது.

PPS முறையில் திரள் மாதிரி தேர்வு :

$$\hat{R} = r_{pps} = \frac{\bar{\bar{y}}_{pps}}{\bar{\bar{x}}_{pps}}$$ என்பதாகவும்

$$V(r_{pps}) = \frac{1}{\bar{\bar{X}}^2}\left[\frac{1}{mN}\sum_{M} N_i(\bar{\bar{Y}}_i - R\,\bar{\bar{X}}_i)^2 + \frac{1}{mN}\cdot\sum_{M}(N_i - n')\cdot\frac{S_i^2}{n'}\right]$$

என்பதாகவும் *PPS* முறை தேர்வில் பெறப்பட்ட முடிவுகளின் அடிப்படையில் பெறப்படுகின்றன.

இதில் $S_i^2 = S_{yi}^2 + R^2 \cdot S_{xi}^2 - 2R \cdot \rho_i\, S_{xi}\, S_{yi}$ ஆகும். $V(r_{pps})$-ன் பிறழ்ச்சியற்ற மதிப்பீடு

$$\hat{V}(r_{pps}) = \frac{1}{\bar{\bar{X}}^2}\left[\frac{1}{m(m-1)}\cdot\sum_{m}(\bar{y}_i - r_{pps}\cdot\bar{x}_i)^2\right]$$

என்பதாகவும் காணலாம்.

பிறழ்ச்சியற்ற விகித மதிப்பீடு (Unbiased Ratio Estimate)

இயல்பாக, பிறழ்ச்சியுடன் அமைகின்ற விகித மதிப்பீட்டினை விட, விகித அடிப்படையில் பிறழ்ச்சியற்ற மதிப்பீடுகள் பெறுவது அதிகமாக கருதப்படுகிறது. அதிக அளவில் படுகைகளையும், சிறு அளவில் அமையும் மாதிரிகளையும் கொண்ட ஆய்வுகளில் இத்தகைய மதிப்பீடுகள் பயன்படும்.

ஹார்ட்லி-ரோஸ் (Hartley-Ross) மதிப்பீடு :

$\frac{y_i}{x_i}$ விகிதங்களின் சராசரியினைக் கருதி, அதில்வரும் பிறழ்ச்சிக்கான திருத்தத்தை இணைத்து, பின்வருமாறு பெறலாம். $\bar{r}$ என்ற மாதிரியின் சராசரியை $\bar{r}$ என்று குறிக்க,

$$E(r) = \bar{r} = \frac{1}{n}\sum_{n} r_i = \frac{1}{n}\sum \frac{y_i}{x_i}$$ என்று வரும்.

எனவே,

$$\frac{1}{N}\sum_{N} r_i (x_i - \bar{X}) = \frac{1}{N}\sum_{N} \frac{y_i}{x_i}\cdot x_i - \left(\frac{1}{N}\sum_{N} r_i\right)\bar{X}$$

$$= \bar{Y} - \bar{X}\cdot E(r_i)$$

$$= \bar{X}[R - E(r_i)]$$

என்பதாக அமையும். மேலும் $E(r_i) = E(\bar{r})$

($\because$ சம வாய்ப்பு முறை)

எனவே, $\bar{r}$-ல் உள்ள பிறழ்ச்சியானது

$$[E(\bar{r}) - R] = -\frac{1}{N.\bar{X}}\cdot\sum_{N} r_i (x_i - \bar{X})$$

என்று பெறப்படுகிறது.

எளிய சம வாய்ப்பு முறையில் அடைந்த முடிவின்படி, $\frac{1}{(N-1)}\sum_{N} r_i (x_i - \bar{X})$ என்பதன் பிறழ்ச்சியற்ற மதிப்பீடு

$$\frac{1}{n-1}\sum_{n} r_i (x_i - \bar{x}) = \frac{n}{n-1}(\bar{y} - \bar{r}\cdot\bar{x})$$

என்று காணப்படும்.

$\therefore$ பிறழ்ச்சிக்காகத் திருத்தப்பட்ட (corrected) மதிப்பீடானது,

$$\bar{r}' = \bar{r} + \frac{n(N-1)}{(n-1)N\cdot\bar{X}}(\bar{y} - \bar{r}\cdot\bar{x})$$

என்றும், இதற்கேற்றபடி அமையும் $\hat{Y}$ மதிப்பீடு பிறழ்சியற்றதாக

$$\bar{r}'\cdot X = \bar{r}\cdot X + \frac{n(N-1)}{(n-1)}\cdot(\bar{y} - \bar{r}\bar{x})$$

என்றும் வருகின்றன.

எடுத்துக்காட்டு :

பின்வரும் பட்டியலிலிருந்து, படுகையின் மொத்ததத்திற்கு மதிப்பீட்டைக் காண்க.

$n = 8$; [படுகையின் மொத்தம் $N = 16$; $X = 106$ என்று கொடுக்கப்படுகிறது]

அலகு எண்	1	2	3	4	5	6	7	8
y_i	8	15	5	7	5	13	11	24
x_i	8	6	1	4	3	5	4	13
r	1·00	2·5	5	1·75	1·67	2·6	2·75	1·85

$$\bar{y} = \frac{\Sigma y}{n} = \frac{88}{8} = 11 \,;$$

$$\bar{x} = \frac{44}{8} = 5{\cdot}5$$

திருத்தப்பட்ட r மதிப்பீடு :

$$r' = \bar{r} + \frac{n\,(N-1)}{(n-1)\,N\,\bar{X}} \cdot (\bar{y} - \bar{r}\cdot\bar{x})$$

இதில்,

$$\bar{r} = \frac{1}{n}\Sigma\frac{y}{x} = \frac{19{\cdot}12}{8} = 2{\cdot}39$$

$$\therefore\; r' = (2{\cdot}39) + \frac{8 \times 15}{16 \times 7 \times \frac{106}{16}}\,[\,11 - (2{\cdot}39 \times 5{\cdot}5)\,]$$

$$= 2{\cdot}39 - 347 = 2{\cdot}043$$

$$\therefore\; \hat{Y} = \bar{r}'\cdot X = (2{\cdot}043) \times 106 = 216{\cdot}6$$

(திருத்தப்பட்டது)

$\hat{Y}$ (திருத்தப்பட்டது)-ன் மாறுபாடானது

V ($\hat{Y}$ திருத்தப்பட்டது) $= V\,(\bar{r}'\cdot X)$

$$= \frac{N^2}{n}(S_y{}^2 + \bar{R}^2 \cdot S_x{}^2 - 2R\rho \cdot S_x \cdot S_y)$$

$$+ \frac{N^2}{n(n-1)} \cdot (S_r{}^2 \cdot S_x{}^2 + S_{rx}{}^2)$$

$\bar{R} = r_i$-ன் முழுமைத் தொகுதி சராசரி ;

$S_r{}^2 = r_i$-ன் முழுமைத் தொகுதி மாறுபாடு.

S_{rx} = முழுமைத்தொகுதியில், (r_i, x_i) இவற்றினிடையே உள்ள உடன் மாறுபாடு (Covariance). இந்த சூத்திரத்தில், *fpc* காரணி கருதப்படுவதில்லை.

n மிகப் பெரிதாக இருப்பின்,

$$V(\bar{r}'X) \simeq \frac{N^2}{n} \sum_{1}^{N} \frac{[\,y_i - \bar{Y} - \bar{R}(x_i - \bar{X})\,]^2}{N-1} \quad \ldots\ldots \quad (1)$$

மேலும்,

$\hat{Y}_r$ மதிப்பீட்டிற்கு,

$$V(Y_r) \simeq \frac{N^2}{n} \cdot \sum_{N} \frac{(y_i - R_{xi})^2}{N-1} \quad \ldots\ldots \quad (2)$$

என்பதாக அமைவது குறிப்பிடத்தக்கது.

(1) (2) இவற்றை ஒப்பிட, பெரிய அளவினதான மாதிரிகளில் $\bar{r}'X$ ஆனது அதிக திட்பத்துடன் அமையவேண்டுமெனில், $[\bar{Y} + \bar{R}(x_i - \bar{X})]$ என்ற கோட்டினில் y_i மதிப்புகள் நெருக்கமாக அமைதல் வேண்டும். இது R_{xi} என்ற கோட்டினை விட அதிக அளவில் இருத்தல் வேண்டும்.

II. மாறிகளின் தொடர்பு மதிப்பீடு : (Regression Estimate)

1. நேர்கோட்டுத் தொடர்பு (Linear Regression)

முன்பு விளக்கப்பட்ட 'விகித மதிப்பீட்டு' முறையினைப் போன்று, தற்சமயம் கருதப்படும் நேர்கோட்டுத் தன்மையில்

(linear) அமையும் 'மாறிகளின் தொடர்பு மதிப்பீடா'னது, y_i என்ற மாறியுடன் தொடர்பு பெற்றிருக்கும் x_i என்ற துணைமாறியினைப் (auxiliary variable) பயன்படுத்தி, மதிப்பீட்டுத் திட்பத்தினை (precision) அதிகரித்தற் பொருட்டு கையாளப்படும் வழிமுறையாகும். நேர்கோட்டு வடிவில் தொடர்பு அமைந்திருப்பினும், அது (0, 0) என்னும் 'ஆரம்பத்'தின் (Origin) வழியாகச் செல்வதில்லை என்பது கவனிக்கத் தக்கதாகும் எனவே, x-ன்மீது அமையும் y-ன் நேர்கோட்டுத் தொடர்பு நிலை அடிப்படையில் மதிப்பீடு நடைபெறும்.

எடுக்கப்படும் மாதிரியின் ஒவ்வோர் உறுப்பிற்கும் (unit) y_i மற்றும் x_i மதிப்புகள் பெறப்படுகின்றன. மேலும் x-ன் முழுமைத் தொகுதியான $\bar{X}$ ஏற்கெனவே கொடுக்கப்பட்டதாகக் கொள்ளப்படும். அந்நிலையில், $\bar{Y}$-ன் நேர்கோட்டுத் தொடர்பு மதிப்பீடு

$$\bar{y}_{LR} = \bar{y} + b(\bar{X} - \bar{x})$$

என அமைகிறது. 'LR' என்பது நேர்கோட்டுத் தொடர்பினையும் (linear regression), 'b' எனும் மதிப்பு, x-ன் மதிப்பு $(x + 1)$ என்று அதிகரிக்கப்படுவதற்கு ஏற்ப y-ல் நிகழும் மாற்றத்தின் மதிப்பீட்டினையும் குறிக்கும். அதாவது, $\bar{x}$ மதிப்பு, சராசரிக்குக் கீழாக இருப்பின், $\bar{y}$ மதிப்பும் $b(\bar{X} - \bar{x})$ என்னும் அளவு (amount) சராசரிக்குக் கீழாக இருக்குமென எதிர்பார்க்கப்படும். முழுமைத் தொகுதி மொத்த மதிப்பீடாக (population total)

$$\hat{Y}_{LR} = N\bar{y}_{LR}$$

என்று கொள்ளப்படும்.

குறிப்பிட்ட தாவர இனத்தைக் கருதி, இலைகளின் எடையைப் பொறுத்து தொடர்பு பெறும் இலைப்பரப்பினைப் பற்றிய ஆய்வினில், இம்மதிப்பீட்டு முறை வாட்ஸன் (Watson) என்பவரால் மேற்கொள்ளப்பட்டது. $\frac{1}{10}$ ஏக்கர் நிலக்கூறுகளைக் கொண்டு அமைந்த முழுமைத் தொகுதியில், உள்ள மரங்களைக் (timber) கணக்கிட, நிலக்கூறுகளின் மாதிரியைத் தெரிவுத்செய்து அதன்மூலம்மொத்த மரங்கள் அளவு (timber volume) அளவிடப் பட்டது. இது யேட்ஸ் (Yates) மேற்கொண்ட ஒரு பரிசோதனையாக அமையக் காணலாம்.

மாறிகளின் தொடர்புப் போக்கு மதிப்பீட்டின் இயல்புகளும், விகித மதிப்பீட்டு இயல்புகளைப் போன்றே அமைகின்றன. தொடர்பு மதிப்பீடானது, பொருத்தமான (consistent) ஒன்றாக அமைகிறது. அதாவது, எடுக்கப்பட்ட மாதிரியானது முழுமைத் தொகுதியாகவே அமைந்திருக்க நேரிடின், $\bar{x} = \bar{X}$ என்று பெறுவதால், தொடர்பு மதிப்பீடானது $\bar{Y}$ என சுருங்குகிறது. மற்றும் இம் மதிப்பீடானது பொதுவாகப் பிறழ்ச்சி உடையது (biased) என்றும், இப் பிறழ்ச்சிக்கும் தரவிலக்கத்திற்கும் இடையில் உள்ள விகிதமானது மாதிரி அளவு அதிகரிக்கப்படுகையில் குறையும் தன்மை உடையது என்றும் காணலாம்.

குறிப்பாக, $b = \frac{\bar{y}}{\bar{x}}$ எனத் தேர்ந்து எடுக்கும் நிலையில்,

$$\bar{y}_{LR} = \bar{y} + \frac{\bar{y}}{\bar{x}}(\bar{X} - \bar{x})$$

$$\therefore \quad \bar{y}_{LR} = \frac{\bar{y}}{\bar{x}} \cdot \bar{X} = \hat{Y}_R$$

என்று பெறப்பட்டு, விகித மதிப்பீட்டு (ratio estimate) நிலை அமைந்திடுவதைக் காண்க.

2. **நிர்ணயிக்கப்பட்ட 'b' மதிப்புடன், தொடர்பு மதிப்பீடு :**
(Regression Estimate — With preassigned 'b')

மாதிரியில் பெறப்படும் முடிவுகளிலிருந்து b மதிப்பு மதிப்பிடப் பட்டாலும், சில நிலைகளில் முன்கூட்டியே b-யின் மதிப்பினைத் தெரிவுசெய்து கொள்ளுதல் சிறந்தது. திரும்பத் திரும்ப அடிக்கடி நடத்தப்படும் ஆய்வுகளில் ஏற்கெனவே பெறப்பட்ட கணக்கீடுகள்மூலம், b-யின் மாதிரி மதிப்புகள் ஓரளவு மாறாத நிலையில் அமையக் காணலாம். அல்லது, x என்பது சமீபக் கணக்கெடுப்பில் பெறப்பட்ட y-ன் மதிப்பாக அறியப்படின், $b = 1$ என்று எடுத்துக் கொள்ளலாம்.

(அ) எளிய ராண்டம் மாதிரிமுறை.

தேற்றம் : 1.

சாதாரண ராண்டம் மாதிரி முறையில் (SRS), b_0 என்பது ஏற்கெனவே நிர்ணயிக்கப்பட்ட மாறிலியாக அமைந்திடும் நிலையில், நேர்கோட்டுத் தொடர்பு மதிப்பீடானது

$$\bar{y}_{LR} = \bar{y} + b_o (\bar{X} - \bar{x}) \quad \ldots\ldots \quad (1)$$

என்று அமைந்து, பிறழ்ச்சியற்றும் இருக்கும். இம் மதிப்பீட்டின் மாதிரிமுறை சார்ந்த மாறுபாடு (sampling variance)

$$V(\bar{y}_{LR}) = \frac{(1-f)}{n}(S_y^2 - 2b_o \cdot S_{yx} + b_o^2 \cdot S_x^2) \quad \ldots\ldots \quad (2)$$

என்று பெறப்படும்.

இதில், $S_y^2 = \sum_{1}^{N} \frac{(Y_i - \bar{Y})^2}{N-1}$; $S_x^2 = \sum_{1}^{N} \frac{(X_i - \bar{X})^2}{N-1}$;

$$\left(f = \frac{n}{N}\right)$$

$$S_{yx} = \frac{\Sigma (Y_i - \bar{Y})(X_i - \bar{X})}{N-1}$$

என்பதாகும்.

(வரம்புற்ற முழுமைத் தொகுதியில் எந்தவிதக் கருதுகோளும் (assumption) தேவையில்லை.

நிறுவுதல் (Proof)

அடிக்கடி நடத்தப்படும் ஆய்வுகளில் பெறப்பட்ட மாறிலியாக b_o அமைவதால்,

$E(\bar{y}_{LR}) = E(\bar{y}) + b_o E(\bar{x} - \bar{X}) = \bar{Y}$ என்றாகிறது.

மேலும், $\bar{y}_{LR}$ என்பது $[y_i - b_o(x_i - \bar{X})]$ என்ற மதிப்புகளின் மாதிரிச் சராசரியாக, $\bar{Y}$ என்னும் முழுமைத் தொகுதிச் சராசரியுடன் அமைகிறது.

$$\therefore \quad V(\bar{y}_{LR}) = \frac{1-f}{n} \sum_{1}^{N} \frac{[(Y_i - \bar{Y}) - b_o (X_i - \bar{X})]^2}{N-1}$$

$$= \frac{1-f}{n} \cdot (S_y^2 - 2b_o \cdot S_{yx} + b_o^2 \cdot S_x^2)$$

இதுவே தேற்றமாகும்.

துணைமுடிவு :

V ($\bar{y}_{LR}$)-ன் பிறழ்ச்சியற்ற மாதிரி மதிப்பீடு

$$V(\bar{y}_{LR}) = \frac{1-f}{n}(s_y^2 - 2b_o\, s_{yx} + b_o^2 \cdot s_x^2) \qquad \ldots\ldots (3)$$

என்று வழக்கம்போல் பெறப்படுகிறது.

தேற்றம் 2 :

$V(\bar{y}_{LR})$ மதிப்பைச் சிறும அளவினதாக (minimum) ஆக்கும் b_0 மதிப்பானது

$$b_0 = B = \frac{S_{yx}}{S_x^2} = \frac{\sum_{1}^{N}(Y_i - \bar{Y})(X_i - \bar{X})}{\sum_{1}^{N}(X_i - \bar{X})^2} \qquad \ldots(4)$$

என்பதாகும். எனவே சிறும அளவு மாறுபாடு (minimum variance)

$$\underset{min}{V}(\bar{y}_{LR}) = \frac{1-f}{n} S_y^2 (1 - \rho^2) \qquad \ldots(5)$$

என்று அமைந்திருக்கும்.

$$\rho = \frac{\frac{1}{N-1}\sum_{1}^{N}(Y_i - \bar{Y})(X_i - \bar{X})}{S_x \cdot S_y}$$ என்று காணப்படும்.

அதாவது, ρ என்பது முழுமைத்தொகுதியில் y, x இவற்றின் இடையில் உள்ள உடன்தொடர்புக் கெழுவாக அறியப்படுகிறது.

நிரூபணம் :

சமன்பாடு (1)-ல் தரப்பட்ட $V(\bar{y}_{LR})$-ல்

$$b_0 = B + d = \frac{S_{yx}}{S_x^2} + d$$

என்று பிரதியிட,

$$V(\bar{y}_{LR}) = \frac{1-f}{n}\left[S_y^2 - 2\,S_{yx} \cdot \left(\frac{S_{yx}}{S_x^2} + d\right) + S_x^2 \cdot \left(\frac{S_{yx}}{S_x^2} + d\right)^2\right]$$

$$= \frac{1-f}{n}\left[\left(S_y{}^2 - \frac{S_{yx}{}^2}{S_x{}^2}\right) + d^2 \cdot S_x{}^2\right] \qquad ...(5)$$

என்று அமைகிறது.

$d = 0$ எனில் இது மிகச் சிறு (minimum) அளவினதாக ஆகிறது. $\rho = \frac{S_{yx}}{S_x \cdot S_y}$ என்பதைப் பயன்படுத்தி,

$$V_{min}(\bar{y}_{LR}) = \frac{1-f}{n} \cdot S_y{}^2 (1 - \rho^2)$$

என்று பெறலாம்.

இதேபோல், திட்ப இழப்பின்றி b_0 ஆனது B-யிலிருந்து விலகும் அளவினைக் கருதலாம்.

(5), (6) இவற்றிலிருந்து

$$V(\bar{y}_{LR}) = \frac{1-f}{n}\left[S_y{}^2 (1-\rho^2) + (b_0 - B^2) \cdot S_x{}^2\right]$$

$$= V_{min}(\bar{y}_{LR})\left[1 + \frac{(b_0 - B)^2 \cdot S_x{}^2}{S_y{}^2 (1-\rho^2)}\right]$$

என்று பெறப்படும்.

ஆனால் $B \cdot S_x = \rho\, S_y$ என்றிருப்பதால்,

$$V(\bar{y}_{LR} = V_{min}(\bar{y}_{LR})\left[1 + \left(\frac{b_0}{B} - 1\right)^2 \cdot \frac{1-\rho^2}{\rho^2}\right] ..(6)$$

என்று எழுதப்படலாம்.

3. b-மதிப்பு மாதிரியின்மூலம் கணக்கிடப்படும் நிலை :

தெரிவு செய்யப்பட்ட மாதிரி அடிப்படையில் b-யானது மதிப்பிடப்படுகையில், சிறந்த மதிப்பீடாக, மீச்சிறு வர்க்க மதிப்பீடான (least squares estimate)

$$b = \frac{\sum_n (y_i - \bar{y})(x_i - \bar{x})}{\sum (x_i - \bar{x})^2}$$

என்பது பயன்படும்.

இங்கு தரப்படும் வழிமுறையில், முழுமைத் தொகுதி தொடர்பானது நேர்கோட்டில் அமைந்திருக்கவேண்டிய நியதி தேவையில்லை. பெரிய அளவினதான மாதிரிகளைப் பொறுத்த அளவில் (large samples) இம்முடிவுகள் பயன் தரும்.

n அளவுள்ள மாதிரிகளில், $(b - B)$ என்ற வேறுபாடானது $\frac{1}{\sqrt{n}}$ என்ற 'படி' (order) அல்லது ஒழுங்கு உடையதாய் அமையும்.

மேலும்,

$$e_x = y_i - \bar{Y} - B\,(x_i - \bar{X})$$

என்பதாகக் கொண்டு,

$$\sum_N e_i\,(X_i - \bar{X}) = \sum_N (Y_i - \bar{Y})(\bar{X}_i - \bar{X}) - B\sum_N (X_i - \bar{X})^2 = 0$$

என்று (B-யின் வரையறைமூலம்) பெறுகிறோம். மேலும்

$$b = \frac{\sum_n y_i\,(x_i - \bar{x})}{\sum_n (x_i - \bar{x})^2}$$

$$= \frac{\sum_n [\bar{Y} + B\,(x_i - \bar{X}) + e_i]\cdot(x_i - \bar{x})}{\sum_n (x_i - \bar{x})^2}$$

அல்லது $b = B + \dfrac{\sum_n e_i\,(x_i - \bar{x})}{\sum_n (x_i - \bar{x})^2}$ ...(7)

என்று காணலாம்.

$\sum_{n} e_i\,(x_i - \bar{x})/(n-1)$ என்பது $\sum_{n} e_i(x_i - \bar{X})/(N-1)$ என்பதன் பிறழ்ச்சியற்ற மதிப்பீடாவதால், பூச்சிய மதிப்பாக ஆகிறது. திருப்பத் திரும்ப எடுக்கப்படும் n-அளவுடைய மாதிரிகளைக் கொண்டு, மாதிரி உடன்மாறுபாடான (sample covariance) $\sum_{n} e_i\,(x_i - \bar{x})/(n-1)$ என்பதன் பரவல், பூச்சிய மதிப்புள்ள சராசரியினைக் (zero mean) கொண்டதாக அமையும். மாதிரிஉடன் மாறுபாட்டின் தரப்பிழையானது (standard error) $\frac{1}{\sqrt{n}}$-ஆம் 'படி'யினைக்(order) கொண்டதாக இருப்பதால்,n அளவு உள்ள மாதிரிகளில், $\Sigma\, e_i\,(x_i - \bar{x})/(n-1)$ என்பதும், $\frac{1}{\sqrt{n}}$-ஆம் 'படி'யுடன் விளங்கும். ஆனால், n-அளவு மாதிரிகளில், $\Sigma\,(x_i - \bar{x})^2/(n-1)$ என்பது (s_x^2) 'ஒன்று' என்ற 'படி'யினைக் கொண்டிருப்பதால், மேற்கண்ட (7)-ன் மூலமாக, $(b - B)$ என்ற வேறுபாடானது $\frac{1}{\sqrt{n}}$ என்னும் 'படி'யினைக் கொண்டு அமைவதாக முடிவு செய்யலாம்.

தேற்றம் 3 :

B-யின் மீச்சிறு வர்க்க (least squares) மதிப்பீடாக b அமைவதாகவும், $\bar{y}_{LR} = \bar{y} + b\,(\bar{X} - \bar{x})$ என்பதாகவும் இருப்பின், n-அளவுள்ள எளிய சம வாய்ப்பு மாதிரியில் (SRS),

$$V\,(\bar{Y}_{LR}) = \frac{1-f}{n} \cdot S_y^2\,(1-\rho^2)$$

என்பதாகும். இதற்கு, $\frac{1}{\sqrt{n}}$-ம் படியில் உள்ள உறுப்புகளை (terms) நீக்கத்தக்க விதத்தில் n-ஆனது பெரியதாக இருத்தல் வேண்டும்.

பின்வருமாறு இதனை நிரூபிக்கலாம் :

e_i-யின் வரையறையிலிருந்து,

$$e = \bar{y} - \quad - B\,(\bar{x} - \bar{X})$$

என்று பெறப்படும்.

எனவே,

$$\bar{y}_{LR} = \bar{Y} + (b - B)(\bar{X} - \bar{x}) + \bar{e} \text{ என ஆகும்.}$$

e_i-யின் முழுமைத் தொகுதி சராசரி பூச்சியமாவதால், $\bar{e}$ என்பது $\frac{1}{\sqrt{n}}$-ஆம் 'படி'யுடன் விளங்கும். ஆனால் $(b - B)$ எனும் வேறுபாடு $\frac{1}{\sqrt{n}}$-ஆம் படியில் விளங்குவதாக நிறுவப்பட்டது. $(\bar{X} - \bar{x})$ வேறுபாடும் $\frac{1}{\sqrt{n}}$-ம் 'படி'யுடன் அமைவதால், $(b - B) \cdot (\bar{X} - \bar{x})$ என்பது $\frac{1}{n}$-ஆம் 'படி'யுடன் இருக்க நேரிடும். $\frac{1}{\sqrt{n}}$-ஆம் படியில் அமையும் உறுப்புகள் நீக்கப்படும் அளவில், இந்தப் பெருக்கற்பலனைக் கருதாது விடலாம்.

எனவே இறுதியில்

$$\bar{y}_{LR} - \bar{Y} \simeq \bar{e} \text{ என்று பெறலாம்.}$$

ஆனால், $E(\bar{e}) = 0$ எனவும், $E(\bar{e}^2)$ ஆனது ஓர் எளிய ராண்டம் மாதிரியில், e_i மதிப்புகளின் சராசரிக்கு$(\bar{e})$கணக்கிடப் பெறும் மாறுபாடாகவும் பெறப்படுவதால் பின்வரும் நிலை காணப்படுகிறது :

$$\therefore \quad V(\bar{y}_{LR}) = \frac{1-f}{n} \cdot S e/_{x}{}^{2}$$

$$\therefore \quad V(\bar{y}_{LR}) = \frac{1-f}{n} \cdot \frac{\sum\limits_{n} e_i^2}{(N-1)}.$$

இதில்,

$$\sum_{N} e_i^2 = \sum_{N} [(Y_i - \bar{Y}) - B(X_i - \bar{X})]^2$$

$$= \sum_{N} (Y_i - \bar{Y})^2 - B^2 \sum (X_i - \bar{X})^2$$

(B-யின் வரையறைப்படி)

ஆனால், $\rho = \frac{S_{yx}}{S_x \cdot S_y} = B \sqrt{\frac{\Sigma (X_i - \bar{X})^2}{\Sigma (Y_i - \bar{Y})^2}}$

எனவே, இதனைப் பிரதியிட,

$$\sum_N e_i^2 = \Sigma (Y_i - \bar{Y})^2 \quad (1 - \rho^2)$$ என்று பெறப்படும்.

∴ இறுதியில்,

$$y\,(\bar{y}_{LR}) = \frac{1-f}{n} \cdot S_y^2\,(1 - \rho^2)$$ என வரும்.

இதன் மாதிரி மதிப்பீடு,

$$V\,(\bar{y}_{LR}) = \frac{1-f}{n\,(n-2)} \sum_1^n [(y_i - \bar{y}) - b\,(x_i - \bar{x})]^2$$

$$= \frac{1-f}{n\,(n-2)} \left\{ \Sigma\,(y_i - \bar{y})^2 - \frac{[\Sigma\,(y_i - \bar{y})\,(x_i - \bar{x})]^2}{\Sigma\,(x_i - \bar{x})^2} \right\}$$

என்று தரப்படும்.

விகித மதிப்பீடு, மற்றும் இயல்பான சராசரி இவற்றோடு ஒப்பீடு :

இத்தகைய ஒப்பீடுகள் நடத்த வழிசெய்யும் வகையில், விகித மற்றும் தொடர்பு மதிப்பீடுகளின் மாறுபாடுகளின் சூத்திரங்கள் பயன் உடையதாக அமையும் வண்ணம் மாதிரி அளவானது பெரிதாக இருப்பது அவசியம்.

$\bar{Y}$ மதிப்பீடுகளின் மாறுபாட்டு அளவுகள் பின்வருமாறு :

தொடர்பு மதிப்பீட்டு முறை :

$$V\,(\bar{y}_{LR}) = \frac{N-n}{Nn} \cdot S_y^2\,(1 - \rho^2)$$

விகித மதிப்பீட்டு முறை :

$$V\,(\bar{y}_r) = \frac{N-n}{Nn} \cdot [S_y^2 + R^2 \cdot S_x^2 - 2RS_{xy}]$$

இயல்பான மதிப்பீட்டு முறை : $V(\bar{y})$

$$= \frac{N-n}{Nn} \cdot S_y^2.$$

$\rho = 0$ என்று இல்லாத நிலையில், தொடர்பு மதிப்பீட்டின் மாறுபாடு, இயல்பான மதிப்பீட்டின் மாறுபாட்டினைவிடச் சிறிய தாக இருக்கிறது.

மேலும், $-\rho^2 \cdot S_y^2 < (R^2 S_x^2 - 2RS_{xy})$

என்றிருந்தால், தொடர்பு மதிப்பீட்டின் மாறுபாடு, விகித மதிப்பீட்டின்மாறுபாட்டினைவிடச்சிறியதாக இருக்கும். அதாவது, $(\rho S_y - RS_x)^2 > 0$ அல்லது $(B - R)^2 > 0$ என்றிருத்தல் வேண்டும்.

$\therefore$ $B = R$ என்றில்லாத நிலையில், தொடர்பு மதிப்பீடு, விகித மதிப்பீட்டினைவிட அதிகம் திட்பமுடையதாக அமைந் திருக்கும். y_i, x_i இவற்றினிடையே தொடர்பு நேர்கோட்டு வடிவினதாய், (0, 0) என்ற புள்ளி வழியே அமையும் நிலையில் இவ்வாறு காணப்படும்.

எடுத்துக்காட்டு :

$S_y^2 = 6409$; $S_{yx} = 4434$; $S_x^2 = 3898$.

$R = 1{\cdot}27$; $\rho = 0{\cdot}89$; $n = 100$; $N = 256$.

$$V(\hat{Y}_{LR}) = \frac{N(N-n)}{n} \cdot S_y^2 (1 - \rho^2)$$

$$= \frac{256 \times 156}{100} \cdot (6409)(1 - 0 \cdot 79)$$

$$= 5{,}45{,}000.$$

[$V(\hat{Y}_r)$, $V(\hat{Y})$ ஆகியவற்றையும் கணக்கிட்டு ஒப்பிடுக.]

பயிற்சி 1

ஒவ்வொரு மரத்திலும் கனிகளின் எடை (x_i), $N = 200$ மரங்களைப் பொறுத்த அளவில் மதிப்பிட்டுப் பெறப்படுகிறது.

மொத்த எடையானது $X = 11{,}600$ lb என அறியப்படும். *SRS* முறையில் கனிகள் மாதிரியாகத் தெரிவு செய்யப்பட்டு, 10 மரங்களைக் கொண்டு பின்வரும் முடிவுகள் பெறப்பட்டன :

மர எண்	1	2	3	4	5	6	7	8	9	10
உண்மை எடை (y)	61	42	50	58	67	45	39	57	71	53
மதிப்பிட்ட எடை (x)	59	47	52	60	67	48	44	58	76	58

உண்மை எடையின் மொத்தமான Y-ன் மதிப்பீடாக,

$$\hat{Y} = N[\,\bar{X} + (\bar{y} - \bar{x})\,]$$

எனக்கொண்டு, மதிப்பீட்டையும் அதன் தரப்பிழை அளவினையும் கணக்கிடுக.

பயிற்சி 2

$N = 5$ என்ற பின்வரும் முழுமைத் தொகுதியில்,

(i) x-ன் மீது அமையும் y-ன் தொடர்பு நேர்கோட்டு வடிவில் அமைகிறது எனவும்,

(ii) *SRS* முறையில் மாதிரி எடுக்கப்படுகையில், நேர்கோட்டு தொடர்பு மதிப்பீடு பிறழ்ச்சியற்று இருக்கிறதெனவும் காட்டுக :

(y, x) இணைகள் : (3, 0) ; (5, 0) ; (8, 2) ; (8, 3) ; (12, 3).

(ஆ) படுகை முறை மாதிரியில் தொடர்பு மதிப்பீடு :

விகித மதிப்பீட்டினைப்போன்று, இருவகையான தொடர்பு மதிப்பீடுகள், படுகை மாதிரி முறையில் அமைகின்றன.

தனித்த முறையில் (separate), ஒவ்வொரு படுகை சராசரிக்கும் (stratum mean) ஒரு தனித்த தொடர்பு மதிப்பீடு கணக்கிடப்படுகிறது. அதாவது,

$$\bar{y}_{LRh} = \bar{y}_h + b_h\,(\bar{X}_h - \bar{x}_h) \qquad \ldots\ldots (1)$$

$$\therefore\ \bar{y}_{LR_s} = \sum_h \frac{N_h}{N}\cdot \bar{y}_{LRh} = \sum_h w_h \cdot \bar{y}_{LRh} \qquad \ldots\ldots (2)$$

$$w_h = \frac{N_h}{N}$$

உண்மையான தொடர்புக்கெழுக்கள் (regression coefficients) B_h மதிப்புகள் படுகைக்குப் படுகை மாறுபடின், இந்த மதிப்பீடு பொருத்தமான ஒன்றாகும்.

இரண்டாவது வகையான 'இணைத்த முறை' (combined) எல்லாப் படுகைகளிலும் B_h மதிப்பு சமமானதாக இருக்கும் நிலையில் பொருத்தமாக அமையும்.

அதற்கு, முதலில்

$$\bar{y}_{st} = \sum_h \frac{N_h}{N} \cdot \bar{y}_h = \sum_h w_h \cdot \bar{y}_h$$

$$\bar{x}_{st} = \sum_h w_h \cdot \bar{x}_h$$

என அறிந்து, அதன்பின்

$$\bar{y}_{LRC} = \bar{y}_{st} + b\,(\bar{X} - \bar{x}_{st})$$

மதிப்பிடப்படும்.

முதலில், இரு மதிப்பீட்டு முறைகளும் b_h, b ஆகியவை முன்னடியாகத் தெரிவுசெய்யப்பட்ட நிலையில் கருதப்பட்டு கணக்கீடு எளிதாகக் காணலாம். $\bar{y}_{LRh}$ ஆனது $\bar{Y}_h$-ன் பிறழ்ச்சியற்ற மதிப்பீடாவதால், $\bar{y}_{LRS}$ ஆனது $\bar{Y}$-ன் பிறழ்ச்சியற்ற மதிப்பீடாகிறது. ஒவ்வொரு படுகையிலும் மாதிரித் தேர்வு தனித்து (independent) அமைவதால்,

$$V(\bar{y}_{LRS}) = \sum_h \frac{w_h^2 \cdot (1 - f_h)}{n_h} \cdot \left(S_{yh}^2 - 2b_h \cdot S_{yxh} + b_h^2 \cdot S_{xh}^2\right) \quad \ldots\ldots (3)$$

என்பதாகும். $b_h = B_h$ எனில், $V(\bar{y}_{LRS})$ ஆனது சிறும அளவினதாகிறது. B_h = உண்மையான தொடர்புக்கெழு (h ஆவது படுகை) சிறும அளவு மாறுபாடு

$$V_{min}(\bar{y}_{LRS}) = \sum_h \frac{w_h^2 \cdot (1 - f_h)}{n_h} \left(S_{yh}^2 - \frac{S_{yxh}^2}{S_{xh}^2}\right) \quad \ldots\ldots (4)$$

என்று தரப்படும்.

b-யின் கொடுக்கப்பட்ட மதிப்புடன், இணைத்த முறையினைக் கருதிடின், $\bar{y}_{LRC}$ ஆனது $\bar{Y}$-ன் பிறழ்ச்சியற்ற மதிப்பீடாவதைக் காணலாம். மேலும், $\bar{y}_{LRC}$ ஆனது $\bar{y}_{hi} + b\,(\bar{X} - x_{hi})$ என்ற மாறியின் படுகை மாறியின் படுகை மாதிரிமுறை மதிப்பீடாக இருப்பதால்,

$$V\,(\bar{y}_{LRC}) = \sum_{h} \frac{w_h^{2}\cdot(1-f_h)}{n_h}.\ (S_{yh}^{2} - 2bS_{yxh} + b^{2}S_{xh}^{2}) \qquad \ldots\ldots \quad (5)$$

என்று பெறப்படும்.

சமன்பாடு (5)-ல் தரப்பட்ட மாறு பாட்டு அளவினை மிகச்சிறு (சிறும) அளவினதாக ஆக்கும் ' b ' யின் மதிப்பு

$$B_c = \sum_{h} \frac{w_h^{2}\,(1-f_h)\,S_{yxh}}{n_h} \bigg/ \sum_{h} \frac{w_h^{2}\,(1-f_h) S_{xh}^{2}}{n_h} \qquad \ldots\ldots \quad (6)$$

என்று பெறப்படுகிறது.

அதாவது, B_c யானது $B_h = \dfrac{S_{yxh}}{S_{xh}^{2}}$ என்ற படுகைத் தொடர்புக் கெழுக்களின் (stratum regression coefficients) எடையிட்ட சராசரியாக அமைகிறது.

பின்வருமாறு

$$a_h = \frac{w_h^{2}\,(1-f_h)}{n_h}.\ S_{xh}^{2}$$

என்று எழுதுகையில்,

$$B = \frac{\Sigma\, a_h\, B_h}{\Sigma\, a_h}$$

என்று ஆவதால், மேற்கண்டவாறு காணப்படுகிறது.

(4), (5) (b க்குப்பதிலாக B_L பிரதியிடப்பட்டு)

ஆகியவற்றிலிருந்து,

$$V_{min}\,(\bar{y}_{LRC}) - V_{min}\,(\bar{y}_{LRS})$$
$$= \Sigma\, a_h\cdot B_h^{2} - (\Sigma\, a_h).\ B_c^{2}$$

$$= \Sigma a_h (B_h - B_c)^2 \qquad \ldots\ldots \quad (7)$$

என்ற முடிவு பெறப்படுகிறது.

இதன்மூலம், B_h மதிப்பானது எல்லாப் படுகைகளிலும் சமமாக இல்லாதிருக்கையில், இணைத்த முறையில் பெறப்படும் மாறுபாட்டினைவிடக் குறைந்ததாகத் தனித்த மதிப்பீட்டின் மாறுபாடானது அமைந்து விளங்கிடக் காணலாம்.

மாதிரியினின்று தொடர்புக்கெழு மதிப்பிடப்படும் நிலை :

தொடர்பு மதிப்பீடுகளில் பயன்படுத்தப்பெறும் b_h மற்று b மதிப்பீடுகள் அதிக திறனுடன் அமையத்தக்க விதத்தில் உருவாக்கப்படும் நிலையினை முன்பே குறிப்பிட்டோம்.

தனித்த மதிப்பீட்டில்,

$$b_h = \frac{\sum_i (y_{hi} - \bar{y}_h)(x_{hi} - \bar{x}_h)}{\sum_i (x_{hi} - \bar{x}_h)^2}$$

என்ற மதிப்பினை, B_h-ன் சிறும வர்க்க (least squares) மதிப்பீடாகக் கொள்ளலாம் எனக் கண்டறிந்தோம்.

ஒவ்வொரு படுகையிலும்,

$$V(\bar{y}_{LRS}) = \sum_h \frac{w_h^2 \cdot (1 - f_h)}{n_h} \cdot S_{yh}^2 (1 - \rho_h^2) \quad \ldots(8)$$

என்று அமைகிறது. இதற்கு, எல்லாப் படுகைகளிலும் n_h பெரிதாக இருத்தல் வேண்டும். [தேற்றம் (3)-ன்படி.] இம்மாறுபாட்டின் மாதிரி மதிப்பீடு,

$$s_{yxh}^2 = \frac{1}{(n_h - 2)} \left[\sum_i (y_{hi} - \bar{y}_h)^2 - b_h^2 \sum_i (x_{hi} - \bar{x}_h)^2 \right]$$

என்ற மதிப்பினை $S_{yh}^2 (1 - \rho_h^2)$ என்பதற்குப் பிரதியிட்டு,

$$v(\bar{y}_{LRS}) = \sum_h \frac{w_h^2 \cdot (1 - f_h)}{n_h (n_h - 2)} \left[\sum_i (y_{hi} - \bar{y}_h)^2 - b_h^2 (x_{hi} - \bar{x}_h)^2 \right]$$

என்பதாகக் கணக்கிடப்படும்.

இந்நிலையிலும், பிறழ்ச்சிக்கும் தரப்பிழைக்கும் இடையில் உள்ள விகிதமானது கணிசமாக அதிகரிக்கும் இடர்ப்பாடு ஏற்படுகிறது. ஒவ்வொரு படுகையிலும், $\bar{y}_{LRh}$ என்ற மதிப்பீடுகள் $\frac{1}{n_h}$ ப்படியில் அமையத்தக்க பிறழ்ச்சிகளுடன் விளங்கும்

எனவே, $\bar{y}_{LRS}$-ன் மொத்தப் பிறழ்ச்சியும் $\frac{1}{n_h}$-ஆம் படியுடனே காணப்படும்.

இணைத்த மதிப்பீட்டின்படி, $b = B_c$ என்னும் நிலையில், மாறுபாடானது மீச்சிறு அளவினதாக அமையும் எனக் கண்டறியப்பட்டது.

எனவே,

$$b_c = \frac{\sum_h \frac{w_h^2 (1 - f_h)}{n_h (n_h - 1)} \sum_i (y_{hi} - \bar{y}_h)(x_{hi} - \bar{x}_h)}{\sum_h \frac{w_h^2 (1 - f_h)}{n_h (n_h - 1)} \sum_i (x_{hi} - \bar{x}_h)^2}$$

என்பதனை B_c-யின் மாதிரி மதிப்பீடாகக் கொள்ளலாம். படுகையாக்குதலானது விகித அளவில் (proportional) அமையும் தருணத்தில், $(n_h - 1)$ என்பதற்குப் பதில் n_h-ஐப் பயன்படுத்தி,

$$b_c' = \frac{\sum_h \sum_i (y_{hi} - \bar{y}_h)(x_{hi} - \bar{x}_h)}{\sum_h \sum_i (x_{hi} - \bar{x}_h)^2}$$

என்று எளிய முறையில் சுருங்கக் காணலாம்.

$$(\bar{y}_{LRC} - \bar{Y}) = \bar{y}_{st} - \bar{Y} + b_c (\bar{X} - \bar{x}_{st})$$

$$= [\bar{y}_{st} - \bar{Y} + B_c (\bar{X} - \bar{x}_{st})]$$

$$+ (b_c - B_c)(\bar{X} - \bar{x}_{st})$$

என்றபடியால், (b_c-யின் மாதிரிப் பிழைகளை நீக்கத்தக்க அளவில்)

$$V(\bar{y}_{LRC}) = \sum_h \frac{w_h^2 (1 - f_h)}{n_h} (S_{yh}^2 - 2B_c S_{yxh} + B_c^2 \cdot S_{xh}^2)$$

என்று பெறப்படும்.

(b_c, b_c' ஆகிய மதிப்பீடுகள் பிறழ்ச்சி உடையனவாகும்). இம்மாறுபாட்டின் மாதிரி மதிப்பீடானது

$$v(\bar{y}_{LRC}) = \sum_h \frac{w_h^2 \cdot (1 - f_h)}{n_h (n_h - 1)} \sum_i [(y_{hi} - \bar{y}_h) - b_c (x_{hi} - \bar{x}_h)]^2$$

என்பதாக அமைகிறது.

ஓர் ஒப்பீடு (A Comparison)

தனித்த முறை அல்லது இணைத்த முறை இவற்றில் எதனை ஒரு குறிப்பிட்ட நிலையில் எடுத்துக்கொண்டு பயன்படுத்துவது என அறிய, திட்டவட்டமான விதியினை உருவாக்க இயலாது. நிலைமைக்கு ஏற்ப, பொருத்தமான முறையினைத் தேர்ந்தெடுக்க வேண்டும்.

தனித்த மதிப்பீட்டு முறையில், ஒவ்வொரு படுகையின் உள்ளும், மாதிரிகள் சிறு அளவில் இருப்பின் பிறழ்ச்சியானது அதிகரிக்கக் காணலாம். தொடர்புக் கெழுக்களில் ஏற்படும் மாதிரிப் பிழைகள் பெரும்பாலும் மதிப்பீட்டின் மாறுபாட்டினைப் பாதிப்பதாக இருக்கும். இதேபோல, இணைத்த மதிப்பீட்டு முறையில், முழுமைத் தொகுதி தொடர்புக் கெழு படுகைக்குப் படுகை வேறுபடுவதாக இருப்பின், அதன் மாறுபாடு பெரும் அளவில் அதிகரிக்கும்.

தொடர்புகள் நேர்கோட்டுத் தன்மையிலும், B_h மதிப்பு எல்லாப் படுகையிலும் சமமானதாகவும் அமைவதாக உறுதிப் படுத்தப்படின், இணைத்த மதிப்பீட்டு முறை பயன்பெறும். ஆனால் நேர்கோட்டுத் தொடர்புடன், B_h ஆனது படுகைக்கு படுகை மாறுவதாக அமைந்திடின், தனித்த மதிப்பீட்டு முறை சிறந்ததாக அமைகிறது.

10. மாதிரி முறை சாராத பிழைகள்
(Non-Sampling Errors)

[அறிமுகம்—சில காரணங்கள்—ஒரு கண்ணோட்டம்—மாதிரி முறை சாராத பிறழ்ச்சி—மாதிரி முறை சாராத மாறுபாடு—சாதாரண ராண்டம் மாதிரி முறையில்—வீத மதிப்பீட்டில்...—செலவுச் சார்பு—பதிலின்மை சார்ந்த பிழை—பிழை நிர்ணயமும் கட்டுப்பாடும்—அளவீட்டுப் பிழைகள்—பதிலளிப்பில் பிழைகள்—முடிவுரை.]

1. அறிமுகம் :

பல்வேறு வகையான மாதிரி முறைகளைப் பற்றிய விரிவான ஆய்வினில், கருதப்படும் முழுமைத் தொகுதியின் ஒவ்வொரு கூறினுக்கும் (unit) உண்மையான ஒரு மதிப்பினைப்பெறமுடியும் என்று எடுத்துக்கொண்டு, எந்தவிதத் தவறும் அல்லது பிழைகளும் இன்றி, பெறும் விவரங்களை நல்லமுறையில் தொகுத்து முடிவுகளை வெளியிடப் பெரும் அளவில் முயற்சி எடுக்கப் படுகிறது எனக் கண்டோம். ஆனால், விவரங்களைக் கண்டறிந்து குறிப்பதிலோ அல்லது ஏதேனும் செயல்முறைப்படுத்தும் நிலையிலோ, பிழைகள் ஏற்படுவது தவிர்க்க இயலாத ஒன்று என்பது கவனத்தில் கொள்ளத்தக்கது. உதாரணமாக பட்டியல் அமைக்கையில் ஏற்படும் தவறுகள் அல்லது பிழைகள், இறுதி முடிவுகளை அதிகம் பாதிக்க வழி ஏற்படுகிறது. எனவே, மாதிரி முறைகளைச் சார்ந்த அளவில், மாதிரி மதிப்பீட்டிலிருந்து முழுமைத் தொகுதி இயல்பைப்பற்றிக் கணிப்பீடு செய்யும் நிலையில் மட்டும் அல்லாது, வேறு பல காரணங்கள் பற்றியும் பிழைகள் ஏற்பட வாய்ப்புண்டு.

மாதிரித் தேர்வு முறைகளைச் சார்ந்து ஒருவகைப் பிழைகள், 'மாதிரி முறைசார்-பிழைகள்' (Sampling Errors), என்றும் இந்தக் காரணத்தினால் அல்லாது, வேறு ஏதேனும் வகையில் ஏற்படும் 'மாதிரி முறை சாராப் பிழைகள்' (Non-Sampling Errors), என்றும் பொதுவாகப் பிழைகள் இருவகையாகப் பிரிக்கப்படுவதை, நூலின் அறிமுகப் பகுதியில் கண்டோம். இதில் இரண்டாவது வகையான 'மாதிரி முறையினைச் சாராப்

பிழைகளின் தன்மையை ஓரளவு இங்கு விரிவாகக் கருதுகிறோம். பெரிய அளவில் ஆய்வு நடைபெறும் நிலையில், இத்தகைய பிழைகள் ஏற்படுவதைத் தடுக்க இயலாது என்பது அனுபவத்தின் அடிப்படையில் பெறப்படும் முடிவாகும். சில சூழ்நிலைகளில், மாதிரி முறையினுல் ஏற்படாத பிழைகளின் அளவு அதிகமாக இருப்பின், தக்க கவனம் செலுத்தப்படவேண்டிய அவசியம் ஏற்படுகிறது.

பொதுவாக முழுக்கணிப்பு ஆய்வுகளில் (Complete enumeration) மாதிரிமுறை சாராத பிழைகளும், மாதிரிக் கணிப்பு ஆய்வுகளில் (sample surveys) மாதிரிமுறை சார்ந்த மற்றும் சாராத இருவகையான பிழைகளும் ஏற்படுவதைக் காணலாம். இவற்றை நன்கு கட்டுப்படுத்தி, ஒரு குறிப்பிட்ட அளவு குறைப்பதற்கேற்ற வழிமுறைகளை அறிந்திடல் வேண்டும். மாதிரி அளவு அதிகரிக்கப்படுகையில் (size), மாதிரிமுறை சார்ந்த பிழையின் அளவுகுறையும் நிலையினை வெவ்வேறு மாதிரித் தேர்வு முறைகளில் விளக்கினேம். இதற்கு மாருக, மாதிரிஅளவு அதிகமாகும்போது, மாதிரிமுறையினைச் சார்ந்திராத பிழையின் அளவும் கணிசமாக அதிகரிக்கும் என்பதனைக் கவனத்தில் கொள்ளவேண்டும்.

பெரிய அளவில் (large-scale) அமைந்த கணிப்புக்களிலும் ஆய்வுகளிலும், புள்ளி விவரங்களைக் கண்டறிதல் மற்றும் அவற்றைத் தொகுத்து அட்டவணைப்படுத்துதல் போன்ற பல்வேறு கட்டங்களிலும் நிகழத்தக்க, மாதிரித்தேர்வு முறைகளை முற்றிலும் சார்ந்திராத பிழைகளை அல்லது தவறுகளை நன்கு மதிப்பிட்டு அறிந்து, இயன்றவரையில் கட்டுப்படுத்தவேண்டியது இன்றியமையாததாகிறது.

எனவே, இத்தகைய தவறுகளை ஓரளவு கட்டுப்படுத்த, மாதிரித் தெரிவுமுறைகள் பெரும் அளவில் உதவிபுரிகின்றன. இத்தகைய பிழைகளை அளவீடு செய்து, கட்டுப்படுத்தி நிலைப் படுத்தும் வழிமுறைகளை விரிவான அளவில் மஹலநோபிஸ் (Mahalanobis), லாஹிரி (Lahiri), டர்பின் (Durbin) போன்ற பலர் ஆராய்ந்து கண்டறிந்திருக்கின்றனர். ஹான்ஸென் (Hansen), சுகாத்மே (Sukhatme) போன்ரோரும் இப்பிரச்சினை யினைப்பற்றி, முழுக்கணிப்பு மற்றும் மாதிரி ஆய்வு இருநிலை களிலும் நன்கு ஆராய்ந்திருக்கின்றனர் என்பதும் குறிப்பிடத் தக்கது. ஆய்வினைத் திரும்பவும் நடத்துவதன்மூலமும், முடிந்த

பின் மீண்டும் சரிபார்க்கும் முறைகளின்மூலமும், மாதிரிமுறை சாராத பல்வேறு பிழைகள் கணிக்க (மதிப்பிட) ப்படுகின்றன.

மாதிரிமுறையினைச் சாராத பிழைகள் எவ்வாறு நிகழ்கின்றன ?

ஒரு ஆய்வின் திட்டநிலையில் தொடங்கி இறுதிமுடிவு அடையும் நிலைவரையிலான பல்வேறு கட்டங்களில் இத்தகைய பிழைகள் ஏற்படுகின்றன. முக்கியமாக ஆய்வுப் பரப்பினைச் சரிவரக் குறிப்பிடாமல் இருத்தல், ஆய்வின் நோக்கத்தைத் தெளிவுற வெளியிடாதிருத்தல், முழுமையாக விவரங்களைச் சேகரிக்காதிருத்தல், தவறான வழிகளைப் பின்பற்றுதல் போன்ற சில காரணங்களால், மாதிரிமுறை சாராத பிழைகள் ஏற்படும் என்று குறிப்பிடலாம்.

பின்வரும் சில காரணங்களை வரையறுக்கிறோம் :

(i) ஆய்வின் அடிப்படை நோக்கத்திற்குப் பொருந்தாத வகையிலும், போதாதநிலையிலும் உள்ள விவரங்களைக்குறித்தல்.

(ii) எல்லைகளில் உள்ள கூறுகளைத் தெளிவாக வரையறை செய்யாததால், சில விவரங்களை விட்டுவிடுவது அல்லது ஒரே விவரத்தை மறுபடியும் கண்டறிந்து குறிப்பது அல்லது தவறான விதிமுறைகளைப் பின்பற்றி விவரங்களைச் சேகரித்தல்.

(iii) போதிய அளவு தெளிவு இல்லாத விவரங்கள் அல்லது பட்டியல்கள் அல்லது விளக்கங்கள் இல்லாமையால், முறையற்ற அல்லது சூழ்நிலைக்கு ஒவ்வாத பேட்டி முறைகளைக் (interview) கையாளுதல்.

(iv) தகுந்த அனுபவமும் பயிற்சியும் இல்லாத ஆய்வாளரைப் பயன்படுத்துதல்.

(v) பதிலளிப்போரின் கவனக் குறைவால் கிடைக்கும் தவறான விவரங்களின் மூலம் விளையும் சிக்கல். பதிலின்மையும் இதில் அடங்கும் பிரச்சினை ஆகும்.

(vi) மேலாளர்களின் தகுந்த கண்காணிப்பும் பரிசீலனையும் இல்லாதிருத்தல்.

(vii) அடிப்படை விவரங்களைச் சரிவர கவனிக்காது விடுதல்.

(viii) செயல்முறைப் படுத்துவதில் ஏற்படும் தவறுகள் அதாவது, குறியிடல் (coding), துளைஅட்டையில் பதிவுசெய்தல் (punch card), சரிபார்த்தல், பட்டியல் அமைத்தல் (Tabulation) போன்ற நிலைகளில் ஏற்படும் பிழைகள்.

(ix) விவரங்களை ஒழுங்குபடுத்தி வெளியிடுதலில் அல்லது விளக்கப்படங்களை அமைப்பதில் நிகழக்கூடிய பல தவறுகள்.

இவை அனைத்தும் பலவகைகளில் தவறுகள் ஏற்பட உதவும் காரணங்களாக, ஓரளவு கருதலாம். மாதிரி ஆய்வில், தக்க படியான மாதிரிக் கூறுகளைத் தெரிவு செய்யாததாலும், தவறான 'சட்ட' (frame) அமைப்பினை எடுத்துக் கொள்வதாலும் மற்ற குறைபாடுகளாலும், இத்தகைய தவறுகள் ஏற்பட வாய்ப்பு உண்டு.

விரிவான முறையில், மாதிரிமுறையினைச் சாராத பிழைகளை மூவகையாகப் பிரிக்கலாம். அவையாவன :

(அ) குறிப்பீட்டுப் பிழைகள் (Specification error)

(ஆ) உறுதிப்படுத்துதல் பற்றியபிழை(Ascertainment error)

(இ) அட்டவணைப்படுத்துதலில்பிழைகள் (Tabulation error)

இம்மூன்றினுள், (அ) - வகையானது ஆய்வின் திட்டமிடும், நிலையிலும், (ஆ) - வகையானது கள ஆய்வுப் பணியிலும், (இ) - வகைப்பிழை இறுதிநிலையான அட்டவணைப்படுத்துதலிலும் காணப்படும் என்று கூறலாம்.

ஆய்வு நோக்கத்துக்குச் சற்றும் பொருந்தாத அல்லது போதிய அளவில் அமையாத விவரங்களைக் குறித்துக் கொள்வதாலும், விவரங்களைத் தெளிவாக வரையறைப்படுத்தாததாலும், தவறான கணிப்புமுறைகளை (enumeration) பயன்படுத்துவதாலும், அல்லது சரியான முறைகள் மூலம் தொடர்பு கொள்ளாததாலும், (அ) வகையான தவறுகள் நிகழுகின்றன. விவரங்கள் தெளிவாகக் குறிக்கப்படுவது ஆய்வின் அடிப்படைத் தேவையாகிறது.

இரண்டாவது வகையில், ஆய்வாளரின் அனுபவக் குறைவு அல்லது போதிய பயிற்சி இல்லாமை பதிலளிப்போரின் (respondents) அலட்சியப்போக்கு அல்லது கவனக்குறைவு, மேலாளரின் (Primary staff) போதிய கண்காணிப்பு இல்லாதிருத்தல் போன்ற காரணங்களினால், விவரங்களை உறுதிப்படுத்தும் நிலையில், (ஆ) போன்ற தவறுகள் ஏற்படுகின்றன.

மூன்றாவதாக, அடிப்படை விவரங்களை நன்கு பரிசீலனை செய்யாமலும், அட்டவணைப்படுத்தும் ஒழுங்குமுறைகளை நன்கு கடைப்பிடிக்காமலும், தவறான முறையில் செயல்படுவதாகவும் இருக்கும் நிலையில், (இ) வகையான பிழைகள் நேரிடுகின்றன.

மேலும், (ஆ) வகையில் குறிப்பிடப்பட்ட தவறுகள் இரு பிரிவுகளாகக் கருதப்படுகின்றன. திரும்பத் திரும்ப ஒரே விவரத்தைக் கண்டறிவதாலும், அல்லது சில விவரங்கள் கணிப்பிலிருந்து விட்டுப்போக நேரிடுவதாலும் (duplication and omission), முழுமைத் தொகுதி அல்லது மாதிரி உறுப்புக்கள், சிலசமயம் தேவைக்கு அதிகமாகவும் சிலசமயம் முற்றுப்பெருமலும் கணிக்கப்படுவதைக் காணலாம். இத்தகைய தவறுகள் அல்லது பிழைகள் 'நிறைவு சார்ந்த பிழைகள்' (coverage error) எனப் படும். அதாவது, எடுத்துக்கொண்ட தொகுதியில் எவ்வளவு தூரம் ஆய்வு முற்றுப்பெறும் என்பதன் அடிப்படையில், கிடைக்கும் விவரங்களைப் பொறுத்து இப்பிழைகள் அமையும். அடுத்து, ஆய்வில் பெற்ற விவரங்களின் 'அடக்கப்பிழைகள்' (content error) ஆகும். அதாவது ஆய்வாளரின் அல்லது பதிலளிப்போரின் தவறுகள் மூலமாக, பிழையுள்ள (தவருன) விவரங்கள் குறிக்கப் பட்டு, அதன் மூலம் இறுதி முடிவுகள் பாதிக்கப்படலாம். இவை அட்டவணைப்படுத்தும் நிலையிலும் தொடர்ந்து வரும் தன்மையன ஆகும் என்பது குறிப்பிடத்தக்கது.

3 மாதிரிமுறை சாராத பிழைகள்பற்றிய ஒரு கண்ணேட்டம் :

இத்தகைய பிழைகளைப் பற்றிக் கணித அடிப்படையில் பரிசீலனை செய்வதற்குமுன் மாதிரி ஆய்வுகளின்மூலம் பெறப் படும் ஒரு மதிப்பீட்டின் மொத்தப் பிழையின் (total error) தன்மையினை சிறிது கருதவேண்டும். பொதுவாக மாதிரிஆய்வில் பெற்ற மதிப்பீட்டு அளவைக்கும், முழுமைத் தொகுதியின் உண்மையான (true) மதிப்பிற்கும் இடையில் இருக்கும் வேறுபாடு, 'மொத்தப் பிழை' அளவினைக் குறிப்பதாகக் கொள்ள லாம். மாதிரிக் கூறுகளை முற்றிலும் சரியாக இனம் பிரித்து அறிந்து விவரங்களைத் தவறின்றிக் குறித்துப் பட்டியல் அமைத் திடும் நிலையில், இந்த மொத்தப் பிழையானது முற்றிலும் மாதிரி முறையினைச் சார்ந்ததாகவே (sampling error) இருக்கும். இதனை, சராசரி வர்க்கப்பிழை (mean square error) மூலமாக மதிப்பிட்டு அறிகிறோம்.

இத்தகைய சராசரி வர்க்கப் பிழையானது, மாதிரிமுறையில் பிறழ்ச்சி மற்றும் மாதிரிமுறை சார்ந்த மாறுபாடு (sampling bias and sampling variance) என்ற இரு பகுதிகளைக் கொண்டதாகும். மாதிரி சாராத பிழைகளும் இருப்பின், மொத்தப் பிழையானது மாதிரிமுறை சார்ந்த மற்றும் சாராத இருவகைப் பிழைகளுடன் அமையும்.

இங்கு, மாதிரிமுறை சாராத பிறழ்ச்சி, மாதிரிமுறை சாராத மாறுபாடு என்ற இரண்டின் மூலமாக மாதிரி சாராத பிழைகளை மதிப்பிட்டுக் கணிக்கிறோம்.

ஆய்வு நடைபெறும் சூழ்நிலைக்கு முற்றிலும் தொடர்பற்ற விதத்தில் ஏதேனும் ஓர் இயல்பின் உண்மையான மதிப்பு அமைந்திருக்கும். ஆனால், ஆய்வில் அறிவிக்கப்படும் மதிப்பானது (reported), அதன் சூழ்நிலையைப் பொறுத்தே அமையும். உதாரணமாக, ஒரு குறிப்பிட்ட காலத்தில் ஒருவரது வயது, வருமானம், அல்லது ஒரு நகரில் உள்ள மக்கட்தொகை, அல்லது ஒருவரது மாதச் செலவு போன்ற விதத்தில் அமைகின்ற உண்மையான மதிப்புகளை எளிதில் வரையறை செய்ய இயலும். ஆனால், சில நிலைகளில் இது எளிதல்ல. அதாவது, ஒரு மனிதரது பொது அறிவுத் தன்மை, சமூக அடிப்படையில் எடுக்கப்படும் நடவடிக்கைபற்றி ஒருவரது தனிக்கருத்து, ஏதேனும் ஒரு பொருள் பற்றிய விருப்பு வெறுப்பு போன்ற நிலைகளில் உண்மையான மதிப்புகளை அறிவது இயலுவதில்லை எனலாம். சூழ்நிலைக்கு ஏற்றவிதத்தில வரையறுக்கப்படும் மதிப்பீடுகளைப் பயன்படுத்தவேண்டிய அவசியம் ஏற்படுகிறது.

ஆய்வின் நியாயமான சூழ்நிலைகட்கு அல்லது ஏற்கத்தக்க சூழ்நிலைகட்குப் பொருந்தும் விதத்தில் வரையறைகளைக் கொண்டு ஏதேனும் ஒரு மாதிரிக் கூறினை எடுத்துக் கொள்வோம். இதனைப் பற்றி ஆய்வு நடத்தும் பொருட்டு ஆய்வாளர்கள் கொண்ட ஒரு மாதிரியும், செயல்முறைப்படுத்தும் அலுவலர்கள் கொண்ட மற்றொரு மாதிரியும் கருதப்படும். வெவ்வேறு பிரிவினரால் தனித்தனியாக ஒரே மாதிரிக் கூறில் ஆய்வு நடத்தப்படும் பொழுது, பலவகையான பிழைகள் ஏற்படுவதால், முடிவுகள் வெவ்வேறு விதத்தில் அமைவதைக் காணலாம். எனவே, உறுப்புகளை (units) அல்லது அலகுகளைத் தெரிவு செய்வது, ஆய்வாளர்களைத் தெரிவு செய்வது, கணக்கிடுவோரைத் தெரிவு செய்வது போன்றமூன்று தெரிவுகளும் வாய்ப்பின் அடிப்படையில் அமைதல் நன்று.

மதிப்பீட்டு அளவையின் (estimator) எதிர்பார்க்கும் மதிப்பிற்கும், உண்மையான மதிப்பிற்கும் இடையில் அமையும் வேறுபாடு, 'மொத்தப் பிறழ்ச்சி' (total bias) எனக் கருதப் பட்டால், இதில் 'மாதிரிமுறை சார்ந்த பிறழ்ச்சி' மற்றும் 'மாதிரி முறை சாராத பிறழ்ச்சி' என்று இருபகுதிகள் வருவதைக்காணலாம். எனவே, பொதுவாக, மொத்தப்பிழையானது, மாதிரிமுறையுடன்

தொடர்புற்ற பிறழ்ச்சி மற்றும் மாறுபாடு, மாதிரிமுறையுடன் சிறிதும் தொடர் பற்ற பிறழ்ச்சி மற்றும் மாறுபாடு, மேலும் மாதிரி மற்றும் மாதிரி முறை சாராப் பிழைகளின் மூலகாரணம் இவற்றின் கூட்டுப் பாதிப்பு (interaction) போன்றவற்றை உள்ளிட்டதாக அமைகின்றது.

4. மாதிரிமுறை சாராத பிறழ்ச்சி (Non-sampling Bias)

மாதிரிக் கூறுகளைத் தெரிவு செய்தல் நிலையிலும், ஆய்விற்குத் தேவைப்படும் பணியாளர்களைத் தெரிவுசெய்தல் நிலையிலும், வாய்ப்பின் அடிப்படை நியதி பின்பற்றப்படுவதாகக் கொள்வோம் (randomisation) வாய்ப்பின் அடிப்படையில் தெரிவு நடைபெறுவதை 'ராண்டமாக்குதல்' என்று குறிப்பிட்டால் எளிதில் புரிந்துகொள்ள இயலும்.

$\bar{Y}$ என்ற முழுமைத் தொகுதியின் சராசரிக்கு, s-என்ற பணியாளர்களை உள்ளிட்ட மாதிரியினின்று, r-என்ற மாதிரி கூறினுக்குப் (sample) பெறப்பட்ட மதிப்பீட்டை $\hat{Y}_{rs}$ என்று கொள்வோம். விவரங்கள் கொண்ட மாதிரிக் கூறினை நிலையாக வைத்து j-யின் மூலம் எதிர்பார்க்கும் மதிப்பினை

$$E_s\,(\hat{Y}_{rs}) = \hat{Y}_r$$

என்று காண்கிறோம். இது $\hat{Y}_r$-ன் மதிப்பீட்டினின்று வேறுபட்டதாகும். இனி, r-இன் மூலம் எடுக்கப்படும் எதிர்ப்பார்க்கப்படும் மதிப்பாவது

$$E_r\,(\hat{Y}_r) = \bar{Y}'$$

என்று அறியப்படும். ஒரு குறிப்பிட்ட ஆய்வுமுறைமூலம் இதற்கு ஒரு பிறழ்ச்சியற்ற மதிப்பீட்டினைப் பெறலாம்.

மேற்கண்ட $\bar{Y}'$ மதிப்பிற்கும், முழுமைத் தொகுதியின் உண்மையான மதிப்பான $\bar{Y}$-க்கும் இடையில் உள்ள வேறுபாடு

$$B_t\,(\hat{Y}_{rs}) = (\bar{Y}' - \bar{Y})$$

என்று கணக்கிடப்பட்டு, 'மொத்தப் பிறழ்ச்சி' (total bias). என்று கூறப்படும்

மாதிரிமுறை சார்ந்த பிறழ்ச்சியானது (sampling bias),

$$B_r\,(\hat{Y}) = E_r\,(\hat{Y}_r) - \bar{Y}$$

என்று தரப்படும். இது உண்மையான மதிப்புக்களைக் கொண்டு பெறப்பட்ட மதிப்பீட்டு எதிர்பார்க்கும் மதிப்பிற்கும் உண்மை

யான தொகுதிச் சராசரி மதிப்பிற்கும் இடையில் காணப்படும் வேறுபாடாக அறியப்படுகிறது.

மேலும் மொத்தப்பிழையானது,மாதிரிமுறை சார்ந்த பிறழ்ச்சி, மற்றும் மாதிரி முறை சாராத பிறழ்ச்சி இவற்றின் கூட்டுத்தொகை எனக் கருதப்பட்டதால்,

∴ மாதிரிமுறை சாராத பிறழ்ச்சி அளவானது

$$B_r(\hat{\bar{Y}}_{rs}) = B_t(\hat{\bar{Y}}_{rs}) - B_r(\hat{\bar{Y}}_r) = \bar{Y}' - B_r(\hat{\bar{Y}}_r)$$

$$\therefore B_s(\hat{\bar{Y}}_{rs}) = E_r(\hat{\bar{Y}}_t - \hat{\bar{Y}}_r)$$

என்று கண்டறியப்படும்.

இது மாதிரி முறையுடன் தொடர்பற்ற விலக்கங்களின் எதிர் பார்க்கும் மதிப்பாகப் பெறப்படுகிறது.

நடைமுறையில் மாதிரி கணிப்பு ஆய்வுகளில், மாதிரித்தேர்வு முறையுடன் தொடர்புற்ற பிறழ்ச்சி அற்ற (unbiased) மதிப்பீடுகளே பயன்படுத்தப்படுவதால், மொத்தப் பிறழ்ச்சி யானது, மாதிரிமுறையோடு தொடர்பற்ற பிறழ்ச்சியினையே உள்ளிட்டதாக அமைகிறது.

இத்தகைய பிறழ்ச்சியினை நன்கு கணக்கிட்டு அறிவதற்குப் பல வழிமுறைகள் வகுக்கப்பட்டிருக்கின்றன. ஆய்வில் பெறப் பட்ட ஏதேனும் புள்ளிவிவரத்தினை, வேறு வழியின்மூலம் மற்றொரு காலவரையில் தனியாகப் பெறப்பட்ட விவரத்துடன் ஒப்பிட்டு, தகுந்த மாற்றங்களை மேற்கொள்ளலாம். இத்தகைய வெளிநிலை (external) சரிபார்த்தல்மூலம், பல்வேறு விவரங்களைப் பற்றிய தெளிவு ஏற்படும். இத்தகைய ஒப்புமையில், ஒரு மூலக் கூறானது மற்றொன்றினைவிட அதிக அளவில் நம்பகத்தன்மை உடையதாக அமைதல் இன்றியமையாதது. அல்லது பெறப் பட்ட விவரங்களிலிருந்து உட்கூறுகளை எடுத்துத் தனித்தனியாக முடிவுகளைப் பெற்றுத் தெளிவு பெறலாம்.

மக்கட்தொகைக் கணிப்புகளில் அல்லது முழுக்கணிப்புகளில் ஆய்வு முடிவுற்றபின் மேற்கொள்ளும் பின்நிலை ஆய்வுகள் மூலம் (post enumeration), சிறந்த வழிமுறைகளைக் கையாண்டு தெளிவான விவரங்களை சரியாகச் அறியலாம். மற்றொரு வழியில், இருவேறு ஆய்வுகளில் பெறப்பட்ட விவரங்களை ஒப்பிட்டு, அவற்றினிடையே இருக்கும் வேறுபாடுகளை நீக்க வழி தேடலாம்.

5 மாதிரிமுறை சாராத மாறுபாட்டு அளவு (Non-sampling variance)

ஏற்கெனவே, மாதிரிமுறைகளில், பல்வேறு மதிப்பீடுகளின் திட்பத்தினை நன்கு அறியும் பொருட்டு அவற்றின், மாதிரிமுறை சார்ந்த மாறுபாடுகளை மதிப்பிட்டு அறிந்தோம். அதற்கு முற்றிலும் மாறாக, இங்கு காணப்படும் மாறுபாட்டு அளவானது மாதிரி முறையுடன் எந்தவிதத்திலும் தொடர்பற்றதாகும்.

முன்பகுதியில் கருதப்பட்ட மதிப்பீடான $\hat{\bar{Y}}_{rs}$—(s-என்ற ஆய்வாளர் - மாதிரியினின்று, r-என்றும் மாதிரிக்கூறினைப் பயன் படுத்திப் பெறப்படும் சராசரி மதிப்பீடு) யின் சராசரி வர்க்கப் பிழை(Mean square error) யானது

$$M(\hat{\bar{Y}}_{rs}) = E_{rs}(\hat{\bar{Y}}_{rs}-\bar{Y})^2$$

என்று வரையறை செய்யப்படும். $\bar{Y}$ என்பது உண்மை மதிப்பு இது மாதிரிசார்ந்த பிழைகளையும் மாதிரி சாராப் பிழைகளையும் கருதிப் பெறப்பட்ட மதிப்பீடாகும். இதில் பிறழ்ச்சி மற்றும் மாடுபாடு, இரண்டும் அமைவதால்,

$$M(\hat{\bar{Y}}_{rs}) = V(\hat{\bar{Y}}_{rs}) + B^2(\bar{Y}_{rs})$$
$$= E(\hat{\bar{Y}}_{rs} - \bar{Y}')^2 + (\bar{Y}' - \bar{Y})^2$$

என்று எழுதலாம். $\bar{Y}'$ என்பது இரு நிலைகளையும் கருதிப் பெறப் படும் மதிப்பீடு என்று ஏற்கெனவே கூறப்பட்டது. ராண்ட மாக்கும் இருநிலைகளையும் கருத்திற்கொண்டு பெறப்பட்ட மாறுபாடாவது :

$$V_{rs}(\hat{\bar{Y}}_{rs}) = V_r E_s(\hat{\bar{Y}}_{rs}) + E_r V_s(\hat{\bar{Y}}_{rs})$$
$$\therefore V_{rs}(\hat{\bar{Y}}_{ij}) = V_r(\hat{\bar{Y}}_r) + E_r E_s(\hat{\bar{Y}}_{rs} - \hat{\bar{Y}}_r)^2$$

என்பதாகக் காண்கிறோம். அதாவது மொத்த மாறுபாடானது மாதிரிமுறையினைச் சார்ந்த மாறுபாடாகவும், சாராத மாறுபாடாக வும் பிரிக்கப்படுகிறது. வலப்புற இரண்டாவது உறுப்பினில்,

$$(\hat{\bar{Y}}_{rs} - \hat{\bar{Y}}_{r\cdot}) = (\hat{\bar{Y}}_{rs} - \hat{\bar{Y}}_{r\cdot} - \hat{\bar{Y}}_{\cdot s} + \hat{\bar{Y}}') + (\hat{\bar{Y}}_{\cdot s} - \hat{\bar{Y}}')$$

என்று பிரித்து எழுதி, அதன்மூலம்

$$E_{rs}(\hat{\bar{Y}}_{rs} - \hat{\bar{Y}})^2 = E_{rs}[\hat{\bar{Y}}_{rs} - \hat{\bar{Y}}_{i\cdot} - \hat{\bar{Y}}_{\cdot s} + \hat{\bar{Y}}']^2$$
$$+ E_s(\hat{\bar{Y}}_{\cdot s} - \hat{\bar{Y}}')^2 \quad \{\hat{\bar{Y}}_{\cdot s} = E_r(\hat{\bar{Y}}_{rs}).$$

என்பதனைப் பெறுகிறோம். இதில், வலப்புற முதல் உறுப்பு மாதிரிமுறை சார்ந்த மற்றும் சாராத இருவகைப் பிழைகளின் கூட்டு வினைவாகவும் (interaction), இரண்டாவது உறுப்பு ஆய்வாளரிடையே உள்ள மாறுபாட்டு அளவாகவும் இருப்பது குறிப்பிடத்தக்கது.

அனைத்து முடிவுகளையும் இணைத்து, மதிப்பீட்டு அளவையின் சராசரி வர்க்கப் பிழையானது [அதாவது $M(\hat{Y}_{rs})$], (i) மாதிரி முறை சார்ந்த மாறுபாடு (ii) மாதிரிமுறை சார்ந்தும், சாராமலும் அமையும் இரு பிழைகளின் கூட்டு வினைவு, (iii) ஆய்வுப் பணியாளரிடை உள்ள மாறுபாடுகள், (iv) மாதிரி முறையினைச் சார்ந்தும், சாராமலும் அமையும் பிறழ்ச்சிகளின் கூட்டு மொத்தத்தின் வர்க்க மதிப்பு ஆகிய நான்கினையும் கொண்டதாக அமைவதை உணரலாம்.

உதாரணமாக, சமவாய்ப்பு முறையில் மாதிரித் தேர்வினைக் கருதுவோம் (SRS முறை) :

6. எளிய ராண்டம் மாதிரி முறையில்

N-அளவுள்ள முழுமைத் தொகுதியிலிருந்து n-அளவுள்ள ஒரு எளிய சமவாய்ப்பு மாதிரியானது எடுக்கப்படுவதாகக் கொள்வோம். திரும்பவைக்கும் முறையைக் கருதுவோம். இந்த n- கூறுகளை, ஒவ்வொன்றும் 'm' கூறுகள் கொண்ட k துணை மாதிரிகளாக (சம அளவில்) பிரிக்கப்படுவதாகக் கொண்டு, அவற்றில் K ஆய்வாளர்களிலிருந்து ராண்டம் முறையில் தெரிந்தெடுக்கப்பட்ட k-ஆய்வாளர்கள் பணிமேற்கொள்வர் என்றும் கொள்வோம்.

Y_{ij} என்பது j-ஆவது ஆய்வாளர் ($j : 1, 2 \ldots k$), i-ஆவது தொகுதிக் கூறினைப்பற்றித் தரும் ($i : 1, 2 \ldots N$) மதிப்பாகக் கருதுவோம். இதன் உண்மை மதிப்பு Y_i என்க.

y_{ij} என்பது மாதிரிக் கூறின் மதிப்பாகும் $(i : 1, 2 \ldots m)$
$j : 1, 2 \ldots k$

எனவே, முழுமைத் தொகுதிச் சராசரியின் மதிப்பீடு

$$\bar{y} = \frac{1}{n} \sum_{k} \sum_{m} y_{ij} \qquad (n = km)$$

என்று பெறப்படுகிறது.

ராண்டமாக்கப்பட்ட அமைப்பினை உடைய இரு நிலைகளையும் கருதிப் பெறப்படும் எதிர்பார்க்கும் மதிப்பாவது :

$$E(\bar{y}) = \frac{1}{N}\sum_{1}^{N} Y_i'$$

என்பதாகும். இதில் $Y_i' = \frac{1}{K}\sum_{j=1}^{K} Y_{ij}$ என்று குறிக்கப்படும்.

மொத்தப் பிறழ்ச்சியானது, பதில் பெறும் விதத்தின் அடிப்படையில் அமைந்து (response),

$$B_t(\bar{y}) = B_s(\bar{y}) = \frac{1}{N}\sum_{1}^{N} (Y_i' - Y_i) = \bar{Y}' - \bar{Y} \qquad \ldots(1)$$

என அமைகிறது.

மதிப்பீட்டின் மாறுபாட்டு அளவு

$$V_{rs}(\bar{y}) = V_r \cdot E_s(\bar{y}) + E_r \cdot V_s(\bar{y})$$

(r, s ஆகியவை ராண்டமாக்கும் நிலைகளைக் குறிக்கும்).

மாதிரிக் கூறுகளை நிலையானதாகக் கொண்டு, இரண்டாம் நிலையில் ராண்டமாக்கப்படுவதைக் கருதி, $\bar{y}$-ன் கட்டுப்பாட்டு எதிர்பார்க்கும் மதிப்பினை

$$E_s(\bar{y}) = \frac{1}{n}\sum_{1}^{n} y_i'$$

என்று பெறுகிறோம். இதில் $y_i' = \frac{1}{K}\sum_{j=1}^{K} y_{ij}$ ஆகும். முதல் நிலையில் ராண்டமாக்கப்படுவதை ஒட்டி, இதனுடைய கட்டுப்படாத எதிர்பார்க்கும் மதிப்பு (unconditional expectation)

$$V_r \cdot E_s(\bar{y}) = \frac{\sigma_r^2}{n} \qquad \ldots(2)$$

என்று பெறப்படும்.

$$\sigma_r^2 = \frac{1}{N} \sum_{i=1}^{N} (Y_i' - \bar{Y}')^2$$

y-ன் கட்டுப்பாட்டு மாறுபாட்டினை (நிலையான மாதிரிக் கூறுகளுடன், இரண்டாம் நிலை ராண்டமாக்கப்படுதலை ஒட்டி)

$$V_s(\bar{y}) = \frac{1}{K} \cdot E_s \left(\frac{1}{m} \sum_{i=1}^{m} y_{ij} - \frac{1}{m} \sum_{1}^{m} y_i' \right)^2$$

$$= \frac{1}{Km^2} \cdot \frac{1}{K} \sum_{j=1}^{K} \left[\sum_{i=1}^{m} (y_{ij} - y_i')^2 \right.$$

$$\left. + \sum_{i=1}^{m} \sum_{i \neq j}^{m} (y_{ij} - y_i') \cdot (y_{i'j}' - y_i') \right]$$

[ஏனெனில், எல்லா j-மதிப்பிற்கும், $V_s \left(\frac{1}{m} \sum_{i=1}^{m} y_{ij} \right)$

ஆனது ஒரே மதிப்புடன் அமைகிறது.]

இனி, முதல்நிலை ராண்டமாக்குதலைக் கருதி, $V_s(\bar{y})$-ன் கட்டுப்படாத மாறுபாட்டினை

$$\frac{1}{km^2} \cdot \frac{1}{K} \sum_{j=1}^{K} \left[\frac{m}{N} \sum_{i=1}^{N} (Y_{ij} - Y_i')^2 \right.$$

$$\left. - \frac{m(m-1)}{N(N-1)} \cdot \sum_{i=1}^{N} \sum_{i' \neq i}^{N} (Y_{ij} - Y') (Y_{i'j}' - Y_i') \right]$$

என்று பெற்று,

$$E_r \cdot V_s(\bar{y}) = \frac{1}{km} \cdot \sigma_d^2 \{1 + (m-1)\rho_w\} \qquad ...(3)$$

என்று இறுதியில் காண்கிறோம்.

$$\text{இதில் } \sigma_d^2 = \frac{1}{kN} \sum_{i}^{N} \sum_{j}^{K} (Y_{ij} - Y_i')^2$$

என்பது, பெறும் பதில்களில் உள்ள விலக்கங்களின் (response) deviations) மாறுபாட்டு அளவாகவும் (தொடர்பற்றவை), $\rho\, w$ என்பது ஒரு ஆய்வாளர் எடுத்துக்கொண்ட மாதிரிக் கூறின் உள்ளிட்ட பதில்களின் விலக்கங்களுக்கிடையில் அமையும் உள்ளீட்டு உடன்தொடர்புக் கெழுவாகவும் (intraclass correlation) அமைகின்றன.

அதாவது

$$\frac{1}{KN(N-1)} \cdot \sum_{j=1}^{K} \sum_{i=1}^{N} \sum_{i' \neq i} (Y_{ij} - Y_i')(Y_{i'j} - Y_{i'})$$

என்று வரும்.

இறுதியாக, $\bar{y}$-ன் மாறுபாடு மற்றும் MSE இரண்டும்

$$V(\bar{y}) = \frac{\sigma_r^2}{n} + \frac{\sigma_r^2}{n} \{1 + \overline{m-1} \cdot \rho_w\} \qquad ...(4)$$

என்றும்

$$M(\bar{y}) = V(\bar{y}) + (\bar{Y}' - \bar{Y})^2 \qquad ...(5)$$

என்றும் அளிக்கப்படுகின்றன.

{ முழுக் கணிப்பில், மாதிரி சார்ந்த மாறுபாடு பூச்சியமாவதால், அதற்கேற்ற மதிப்பான $\bar{y}'$-க்கு,

$$V(\bar{y}') = \frac{\sigma_d^2}{N} \{1 + \overline{m-1}\, \rho_w\} \text{ எனவும்,}$$

$$M(\bar{y}') = V(\bar{y}') + (\bar{Y}' - \bar{Y})^2 \text{ எனவும்}$$

வரையறுக்கப்படுகின்றன. }

(4)-ல் தரப்படும் $V(y)$ ஆனது, மொத்த மாறுபாட்டினில் இடம்பெறும் 'பதில்களிடை மாறுபாடு' என்ற பகுதியையும், இந்த மாறுபாட்டின்மீது ஆய்வாளரின் உள்ளீட்டு உடன் தொடர்பின் ஆதிக்கத்தினையும் எடுத்துக்காட்டுகிறது. பிரிவு உள்ளீட்டு உடன்தொடர்பு நேர்க் கணியமாக, ஒரு குறிப்பிட்ட ஆய்வாளரைக் கருதி, வெவ்வேறு கூறுகளின் பதிலளிப்பு விலக்கங்கள் ஒரே திசை நோக்கிச் செல்வனவாக அமையும். ஆனால், அதிக அளவில் கூறுகளை ஆய்கையில் (m-பெரிய அளவாக), இத்தகைய பதிலளிப்பு மாறுபாடுகள் அதிக அளவினதாக அமைகின்றன.

$V(\bar{y})$-ன் பிறழ்ச்சியற்ற மதிப்பீட்டினை

$$v(\bar{y}) = \frac{1}{k(k-1)} \cdot \sum_{j=1}^{k} (\bar{y}_{\cdot j} - \bar{y})^2 \qquad \ldots(6)$$

$$\left[\bar{y}_{\cdot j} = \frac{1}{m} \sum_{1}^{m} y_{ij} \right]$$

என்று வரையறுக்கலாம்.

அதாவது, ஆய்வாளர்களின் பெரிய தொகுதியினின்று, சம வாய்ப்பின் அடிப்படையில் (இதை ராண்டமாக்குதல் என்கிறோம்.) K ஆய்வாளரைத் தெரிவு செய்து, அவர்கள் k-தொடர்பற்ற மாதிரிகளை ஆய்வு செய்யும் நிலையில், மொத்த மாறுபாட்டிற்கு ஒரு பிறழ்ச்சியற்ற மதிப்பீட்டினைப் பெற இயலும் என்பதாகும்.

ஆய்வாளரிடை உள்ள மாறுபாடானது

$$\sigma_s^2 = \frac{1}{K} \sum_{j=1}^{K} (\bar{Y}'_{\cdot j} - \bar{Y}')^2 \simeq \sigma_d^2 \cdot \rho_w$$

ஆகும்.

ஏனெனில்,

$$\sigma_s^2 = \frac{1}{K} \sum_{j=1}^{K} \left[\frac{1}{N} \sum_{i=1}^{N} (Y_{ij} - Y'_i) \right]^2$$

$$= \frac{1}{KN^2} \sum_{j=1}^{K} \sum_{i=1}^{N} (Y_{ij} - Y_i')^2$$

$$+ \frac{1}{KN^2} \cdot \sum_{j=1}^{K} \sum_{\substack{i=1 \\ i \neq i'}}^{N} (Y_{ij} - Y_i') \cdot (Y_{i'j} - Y_i')$$

$$= \frac{\sigma_d^2}{N} + \frac{N-1}{N} \cdot \sigma_d^2 \cdot \rho_w$$

$\therefore \quad \sigma_s^2 \simeq \sigma_d^2 \cdot \rho_w$ ஆகிறது. (N பெரிதாக)

இந்நிலையில்,

$$\hat{\sigma}_s^2 = k \cdot V(\bar{y}) - \frac{1}{km(m-1)} \sum_{j=1}^{k} \sum_{i=1}^{m} (y_{ij} - \bar{y}_{\cdot j})^2$$

என்பது σ_s^2-ன் பிறழ்ச்சியற்ற மதிப்பீடாகும்.

{ $\because \quad E(\hat{\sigma}_s^2) = \sigma_s^2$ என்று காட்டலாம். }

7. வீத மதிப்பீடு (Estimation of Proportion)

முழுமைத் தொகுதியின் வீத அளவினை மதிப்பீடு செய்யுமிடத்து, பதில் பெறுவதில் ஏற்படும் மாறுபாட்டு அளவைக் (Response varaince) கருதுவோம்.

$Y_{ij} = 1$ எனில், i-ஆவது தொகுதிக்கூறு, j-ஆவது ஆய்வாளர் அறிவிப்பதற்கேற்ப, ஒரு குறிப்பிட்ட பிரிவில் (class) அமைவதாகவும்,

$Y_{ij} = 0$ எனில், அக் கூருனது குறிப்பிட்ட பிரிவில் அமையவில்லை என்றும் வரையறுக்கிறோம்.

P_i' என்பதை i-ஆவது தொகுதிக் கூறுனது (population unit) குறிப்பிட்ட பிரிவில் அமைவதாகத் தெரிவிக்கும் ஆய்வாளர் வீதமாக கொள்க.

N அளவுள்ள தொகுதியிலிருந்து, n-அளவு மாதிரி எடுக்கப்படுகிறது. இதில் ஆய்வு நடத்தும் பொருட்டு, K ஆய்வாளர்கள்

அடங்கிய தொகுதியிலிருந்து, சமவாய்ப்பு அடிப்படையில், k ஆய்வாளர் கொண்ட மாதிரி தெரிவு செய்யப்படுகிறது. அந்நிலையில்

$$\hat{P} = \frac{1}{mk} \sum_{j=1}^{k} \sum_{i=1}^{m} y_{ij} \qquad (n = mk)$$

என்று வீத மதிப்பீடு அமையக் காணலாம்.

இருநிலை கருதிய எதிர்பார்க்கும் மதிப்பு

$$(\hat{P}) = \frac{1}{N} \sum_{1}^{N} P_i' = P'$$

என்பதைப் பெறலாம்.

இதில், விவரங்களை உறுதிப்படுத்துவதில் ஏற்படக் கூடிய பிறழ்ச்சி மட்டுமே காணப்படும், அது $(P'-P)$ என்று கணக்கிடப்படுகிறது.

மேலும்,

$$\sigma_r^2 = \frac{1}{N} \sum_{1}^{N} (P_i'-P)^2$$

எனவும்,

$$\sigma_d^2 = \frac{1}{N} \sum_{1}^{N} P_i' \, Q_i'$$

எனவும் வருவதைக் காணலாம். $(Q_i' = 1-P_i')$.

P-மதிப்பீட்டின் மாறுபாடானது

$$V(\hat{P}) = \frac{1}{N.n} \sum_{1}^{N} (P_i'-P')^2$$

$$+ \frac{1}{Nn} \sum_{1}^{N} P_i' \, Q_i' \{1 + (m-1)\rho_w\}$$

என்பதைக் காண்க.

மொத்த மாறுபாட்டின் பிழழ்ச்சியற்ற மதிப்பீடு

$$v\ (\hat{P}) = \frac{1}{k(k-1)} \sum_{1}^{k} (p_j - p)^2$$

((6)-ல் பெறப்பட்ட அடிப்படையில்)

$p_j = j$-ஆவது ஆய்வாளருக்கு ஒதுக்கப்பட்ட மாதிரிக் கூறின் அடிப்படையில் பெறப்பட்ட மாதிரி வீதஅளவு, $p=$ மொத்தமாகக் கணக்கிடப்பட்ட மாதிரிவீத அளவு.

N அளவு பெரிதாக இருப்பின், ஆய்வாளரின் இடைப்பட்ட மாறுபாடானது

$$\hat{\sigma}_s^2 = k.v\ (\hat{P}) = \frac{1}{k\ (m-1)} \sum_{1}^{k} p_{\cdot j}\, q_{\cdot j}\ [q_{\cdot j} = 1 - p_{\cdot j}]$$

என்று பிறழ்ச்சியற்று மதிப்பிடப்பெறும்.

பிரிவு உள்ளீட்டு உடன் தொடர்பு பூச்சியமாக அமைந்திடின்,

$$V\ (\hat{P}) = \frac{P'\ Q'}{n}$$

என்று சுருங்குகிறது. இது மாதிரி வீத அளவீட்டின் (sample proportion) மாதிரி சார்ந்த மாறுபாடாகும். இதில் மாதிரி முறையுடன் தொடர்புடைய மாறுபாடு மட்டுமன்றி, ஒன்றோடொன்று தொடர்பற்ற பதிலளிப்பு மாறுபாட்டினையும் காணலாம் என்று பெறப்படும்.

$$v\ (\hat{P}) = \frac{pq}{n-1}$$

என்பது, $V\ (\hat{P})$யின் பிறழ்ச்சியற்ற மதிப்பீடாகும்.

{$\because$ $E\ (pq) = (n-1)\ V\ (p)$ ஆகும்.}

8 செலவுச் சார்பு (Cost function)

கொடுக்கப்பட்ட நிலையான செலவிற்கு ஏற்ப, மொத்த மாறுபாட்டினைச் சிறும அளவினதாக்கத்தக்க K (ஆய்வாளர்

எண்ணிக்கை) மற்றும் m (ஒவ்வொரு ஆய்வாளருக்கும் ஒதுக்கப் படும் மாதிரியில் உள்ள கூறுகளின் அளவு) இவற்றின் உத்தம மதிப்புகளைக் காண்பது அடுத்த பிரச்சினை ஆகும்.

செலவுச் சார்பினை

$$C = c_0 + k.c_1 + nc_2$$

என்று கருதுவோம்.

c_0 = மாறாச் செலவு (overhead cost)

c_1 = ஆய்வாளரின் பயிற்சிக்கு சராசரியாக ஆகும் செலவு வீதம்.

c_2 = ஒரு கூறினைப் (unit) பற்றி ஆய்வதற்கு ஆகும் சராசரிச் செலவு. ($n = mk$) என்று கொள்க.)

மொத்த மாறுபாடானது

$$V(\bar{y}) = \frac{\sigma^2 - \sigma_s^2}{n} + \frac{\sigma_s^2}{k}$$

என்று அமையும்.

$$\{\because \sigma^2 = \rho_w \sigma_d^2; \quad \sigma_s^2 = \sigma_r^2 \sigma_d^2\}$$

இம் மாறுபாட்டினைச் சிறும அளவினதாக ஆக்கும்பொருட்டு n மற்றும் k இவற்றின் மூலமாக, கொடுக்கப்பட்ட செலவில்,

$$k = \frac{(c - c_0)}{\sqrt{c_1 \sigma_s^2} + \sqrt{c_2(\sigma^2 - \sigma_s^2)}} \cdot \sqrt{\frac{\sigma_s^2}{c_1}}$$

என்றும்,

$$m = \sqrt{\frac{c_1}{c_2}} \cdot \sqrt{\frac{\sigma^2 - \sigma_s^2}{\sigma_s^2}}$$

என்றும் நிர்ணயிக்கப்படுகின்றன.

9. பதிலின்மை சார்ந்த பிழை (Non-response error)

தெரிவு செய்யப்படும் மாதிரிக்கூறுகளில் சிலவற்றைப்பற்றி விவரங்கள் பெறாமல் போகும் நிலை 'பதிலின்மை' என்று குறிப்பிடப்படும். கருதப்படும் கூறுகளில், முழு விவரங்களும் அறியப்படாத நிலையில், முழு ஆய்வுகளிலும், மாதிரி ஆய்வு களிலும் பிழைகள் நேரிடுகின்றன. பதில் பெறாத நிலையை ஒரு

வகையான மாதிரித் தேர்வுப் பிறழ்ச்சி என்று கொள்ளலாம். இத்தகைய கூறுகள், தெரிவில் இடம்பெறப் பூச்சிய அளவு வாய்ப்பையே பெறுகின்றன. சில விவரங்கள் ஆய்வுநிலையில் விட்டுப்போக நேரிடின், இறுதி முடிவுகள் பிறழ்ச்சியுடன் அமைகின்றன. பதிலின்மைக் கூறுகளின் இயல்புகள், மற்ற வற்றினின்று மாறுபடுவதால் இவ்வாறு நிகழ்கிறது.

'பதிலின்மை' நிலைக்கு ஒரு சில காரணங்களைத் தொகுத்துப் பின்வருமாறு அளிக்கலாம் :

(i) கணக்கெடுப்பில் இடம் பெறாத நிலை (Non-coverage)

சில கூறுகளை அல்லது நபர்களை நேரடியாகச்சென்று ஆய்வு நடத்த முடியாது போகலாம். சிலரது இருப்பிடங்கள் சரிவரத் தெரியாத நிலையில் இவ்வாறு சிலகூறுகள் விட்டுப்போக நேரிடும். கருதப்படும் விவரப்பட்டியல்கள் முழுமை பெற்றிருத்தல் அவசிய மாகிறது. சில இடங்களுக்கு செல்வதில், சரியான போக்குவரத்து வசதி இல்லாததாலும், மாறுபடும் தட்ப வெப்ப நிலைகளாலும், இடர்ப்பாடுகள் ஏற்படுகின்றன. எனவே, அவ்விவரங்கள் விட்டுப்போக வாய்ப்பு உண்டு.

(ii) 'இருப்பிடத்தில்' இல்லாதோர் (Not 'at homes')

பட்டியலில் கண்டபடி வீடுகளில் வசிப்பதாகத் தெரியவரும் சிலர், ஆய்வுக் காலத்தில், தற்காலிகமாக அவ்விடத்தில் இல்லாமல் போகலாம். எனவே, விவரம் சேகரிப்பது பெரும் பாடாக அமையும். உதாரணமாக, கணவனும் மனைவியும் என்று இருவரைமட்டுமே கொண்ட குடும்பங்களில், இருவரும் பணி புரிபவராக அமைந்திடின், அவர்களை வீடுகளில் காண்பது சிரமமாக அமைகிறது. இதன்மூலம் விவரங்கள் அறியப்படாமல் போகலாம்.

(iii) விருப்பமின்மை :

சில கேள்விகட்கு, பதிலளிப்போர் பதிலளிக்க விரும்பாமல் போகும் பட்சத்தில் ஏதேனும் ஒருவகையில் சாமர்த்தியமாக விவரங்களைப் பெறுவது ஆய்வாளர் திறமையையே பொறுத்து அமையும். சில கேள்விகட்கு உடனே பதில் தரமுடியாமலும் போகலாம்.

(iv) மறுப்பு :

சில நபர்கள் பேட்டிக்கே உடன்படாமல் அலட்சியத்துடன், வெறுப்பினையும் தெரிவிப்பதாக நேரிடும் நிலைகளில், விவரங்கள்

கிடைப்பது பிரம்மப் பிரயத்தனமாக அமைகிறது. இந்நிலையில் ஏற்படும் பிறழ்ச்சியைத் தடுப்பதும் கடினமாகும்.

மேற்கூறப்பட்டவை ஒரு சில காரணங்களே ஆகும். மேலும் சிலவற்றிலும் பதிலின்மை ஏற்படக் காரணம் உண்டு. இதன்மூலம் ஆய்வில் பல்வேறு கட்டங்களில் தவறுகள் நேரிடுகின்றன.

இத்தகைய தவறுகள்' பதிலின்மை சார்ந்த பிழை' (Non-response error) என கருதப்படும். இத்தகைய பிழைகள், ஒரு தொகுதியை அல்லது மாதிரியைச் சார்ந்த ஏதேனும் சில கூறுகளைப் பற்றிய விவரங்கள் விட்டுப்போக (omitted) நேரிடுவதால் ஏற்படுகின்றன எனலாம்.

இத்தகைய 'பதில்-கிடைக்காத' கூறுகளின் (அல்லது நபர்களின்) இயல்புகள், பதில் சரிவர கிடைக்கப்பெறும் கூறுகளினின்று எந்த அளவிலும் மாறுபடாத நிலையில், இத்தகைய பிழையினை அவ்வளவாகப் பொருட்படுத்தவேண்டியதில்லை. ஆனால் இது நடைமுறைக்கு ஒவ்வாதது. உதாரணமாக, கிராமப்புற ஆய்வுகளில் வினாப்பட்டியலானது அஞ்சல்மூலம் அனுப்பப்பெறுகையில், சிறு அளவில் நிலங்கள் உடைய விவசாயிகளிடமிருந்து அதிக அளவில் பதில்கள் சரிவரக் கிடைப்பதில்லை. ஆதலால், பதிலின்மையின் அளவும் அதிகரிக்கிறது. திரும்பத் திரும்ப அவர்களைச் சென்று காண்பதன்மூலம் அல்லது வேறு வழிகளில் இத்தகைய பிழைகளை ஓரளவு நீக்க இயலும். முற்றிலும் நீக்குவதென்பது இயலாத காரியமாகும்.

முதல் முயற்சியில் பதில் அளிக்காத நபர்களிடமிருந்து, துணை மாதிரிகளை அமைத்துக்கொள்வதன்மூலம், விவரங்களைப் பெற முயற்சி செய்யலாம். இம்முறையானது ஹான்ஸென் மற்றும் ஹர்விட்ஸ் (Hansen & Hurwitz) இவர்களால் வகுக்கப்பட்டதாகும் (1946).

N அளவு முழுத் தொகுதியினின்று n-அளவு மாதிரி எடுத்து, (திரும்பவைக்காத சமவாய்ப்பு முறை) n_1-உறுப்புகள் பதில் பெறுவனவாகவும், $n_2 \ (= n - n_1)$ உறுப்புகள் (units) பதில் இல்லாதன வாகவும் (nonresponse) கொள்வோம். அடுத்த முயற்சியில், n_2-பதில் இல்லாத கூறுகளைக் கருதி அவற்றினின்று n_2' கூறுகள் (உறுப்புகள்) உள்ள துணைமாதிரியினை (subsample) எடுப்பதின்மூலம், விவரம் சேகரிப்பதில் தனியான கவனம் செலுத்தப்படுகிறது.

$\bar{y}_1$, $\bar{y}_2'$ என்பவற்றை, முதல் முயற்சியில் பதில் பெறப்படும் n_1-உறுப்புகளின் மாதிரிச் சராசரியாகவும், துணை மாதிரியில் உள்ள n_2'-உறுப்புகளின் மாதிரிச் சராசரியாகவும் முறையே கொள்வோம்.

$\bar{Y}$-ன் பிறழ்ச்சியற்ற மதிப்பீடானது

$$\hat{\bar{Y}} = \frac{1}{n}(n_1 \bar{y}_1 + n_2 \bar{y}_2')$$

என்று பெறப்படுவதைக் காண்கிறோம். இதில், ராண்டமாக்கப் படும் அமைப்பு இரு முகங்களில் (two-phases) கருதப்படுகிறது. முதலில், n_1 உறுப்புகளை தெரிவு செய்தல்; இரண்டாவதாக, n_2-பதில் பெறப்படாத உறுப்புகளினின்று அல்லது நபர்களிட மிருந்து, n_2' உறுப்புகளைத் தெரிவுசெய்வது என்ற விதமாக அமைகின்றது. இவ்விரண்டினையும் கருதிப் பெறப்படும், மதிப்பீட்டின் மாறுபாட்டு அளவினை

$$V_{12}(\hat{\bar{Y}}) = V_1 E_2(\hat{\bar{Y}}) + E_1 V_2(\hat{\bar{Y}})$$

என்று குறிப்பிடுகிறோம்.

இரண்டாம்படியாக ராண்டமாக்கப்படும் நிலையில், கட்டுப் பாட்டு எதிர்பார்க்கும் மதிப்பு (conditional expected value)

$$E_2(\hat{\bar{Y}}) = \frac{1}{n}(n_1 \bar{y}_1 + n_2 \bar{y}_2)$$

$$= N\bar{y} \quad \left(\because \quad \bar{y}_2 = \sum_{n_2} y_{2i}\right)$$

எனவும், மாறுபாடானது

$$V_2(\hat{\bar{Y}}) = \frac{1}{n^2}\{n_2(n_2 - n_2')\} \cdot \frac{s_2^2}{n_2'}$$

எனவும் பெறப்படும்.

$$\text{இதில் } s_2^2 = \frac{1}{(n_2 - 1)} \sum_{n_2} (y_{2i} - \bar{y}_2)^2.$$

$\bar{y}$ என்பது n உறுப்புகளையும் (கூறுகளையும்) கருதிப் பெறப்பட்ட மாதிரிச் சராசரியாகவும், y_{2i} என்பது அந்த மாதிரியில் உள்ள

i-ஆவது பதில் அளிக்காத உறுப்பின் மதிப்பாகவும் காணப் படும்.

எனவே,

$$V(\hat{Y}) = V_1(\bar{y}) + E_1 \left\{ \frac{1}{n} \left(\frac{n_2}{n} \right) (k-1) s_2{}^2 \right\}$$

என்று அமையும். $\left(k = \frac{n_2}{n_2'} \right)$.

மேலும், $\frac{n_2}{n}$-ன் எதிர்பார்க்கும் மதிப்பு $\frac{N_2}{N}$ என்றும் n_2 கொடுக்கப்பட்ட நிலையில் காணப்படும் $s_2{}^2$-ன் கட்டுப்பாட்டு எதிர்பார்க்கும் மதிப்பு

$N_2 \cdot \sigma_b{}^2/(N_2 - 1)$ எனவும் அமைகின்றன.

[$\sigma_b{}^2$ = முதல் முயற்சியில் 'பதில் கிடைக்கப்பெறாத' முழுமைத் தொகுதிக் கூறுகளின் இடைப்பட்ட மாறுபாடு.

N_2 = இத்தகைய 'பதில் கிடைக்காத' முழுமைத் தொகுதி கூறுகளின் எண்ணிக்கை.]

∴ மாறுபாடானது

$$V(\hat{Y}) = \frac{N-n}{N-1} \cdot \frac{\sigma^2}{n} + \frac{1}{Nn} (k-1) \cdot \frac{N_2{}^2 \cdot \sigma_b{}^2}{(N_2-1)}$$

என்பதாக முடிவு செய்யப்படும்.

($\sigma^2 = N$ அளவுள்ள முழுமைத் தொகுதியின் மாறுபாடு)

இதற்கு ஏற்ற செலவுச் சார்பு

$$c = c_1 \cdot n + c_2 \cdot n \cdot P + c_3 \cdot \frac{n}{k} \cdot Q$$

என்று கருதப்படலாம்.

c_1 = விவரங்களைச் சேகரிக்கும் முதல் முயற்சியில் செலவு வீதம் ;

c_2 = அட்டவணைப்படுத்த செலவு வீதம் (cost per unit) ;

c_3 = பதில் அளிக்காத கூறுகளினின்று மாதிரி தெரிவு செய்ய ஏற்படும் செலவு வீதம் ;

P : முதல் முயற்சியிலேயே சரியாக பதில் அளிக்கும் கூறுகள் (நபர்கள்) விகிதம் (முழுமைத் தொகுதியில்).

$Q = 1 - P.$

இந்த மொத்தச் செலவினைக் குறைக்க, மதிப்பீட்டு மாறுபாடு கொடுக்கப்படும் நிலையில், n மற்றும் k இவற்றின் உத்தம மதிப்புகள்

$$n = n' \left\{1 + (k-1)\, Q^2 \cdot \frac{(N-1)\,\sigma_b^2}{(N_2 - 1)\,\sigma^2}\right\}$$

$$k = \sqrt{\left\{\frac{N^2 (N_2 - 1)}{N_2^2 (N-1)} \cdot \frac{\sigma^2}{\sigma^{\;2}} - 1\right\} \cdot \frac{c_3 \cdot Q}{c_1 + c_2 \cdot P}}$$

என்பதாகக் கணக்கிடப்படுகின்றன.

இதில் $n' = \dfrac{N\sigma^2}{[\sigma^2 + (N-1) \cdot V]}$ என்ற மதிப்பு, முழு அளவில் பதில் பெறும் (அதாவது விவரங்களைப் பெறும்) நிலைக்கு ஏற்ற V-என்ற மாறுபாட்டின் மதிப்பினை உறுதி செய்யத்தக்க மாதிரி அளவாக அமைகிறது.

$$\sigma^2 = \sigma_b^2 \text{ என்றும், } \frac{N}{N-1} = \frac{N_2}{N_2 - 1} \simeq 1$$

என்று கொண்டால், உத்தம மதிப்பீடுகள்

$$n = n' \{ 1 + (k-1)\, Q \}$$

$$k = \sqrt{\frac{c_3\, P}{c_1 + c_2 \cdot P}}$$

என்றபடி பெறப்படுகின்றன.

10. பிழை நிர்ணயமும், கட்டுப்பாடும் (Measurement and Control)

மாதிரி முறையுடன் தொடர்பற்ற பிழைகளின் அளவை நிர்ணயிக்கக் கையாளப்படும் வழிமுறைகளையும், அவற்றைக் கட்டுப்படுத்த மேற்கொள்ளும் வழிகளையும்பற்றி, மாதிரி அல்லது முழு கணிப்பு ஆய்வினைத் துவங்கும்முன்னரே, தீர யோசித்து முடிவு செய்துகொள்வது நன்று. ஏனெனில், சிலமுறைகள் ஆய்வின் திட்டநிலையிலே (design) அதிகம் தேவைப்படுகின்றன. இதோடு மட்டுமன்றி, இறுதி முடிவுகளில் இடம்பெறும் மாதிரி

முறை சாராத பிழைகளை நன்கு மதிப்பிட்டு அறியவும் தனியான முறைகளை வகுக்கவேண்டியிருக்கும். இதற்காக, விவரங்களைச் சேகரிப்பதிலும், தொகுத்து ஆய்வதிலும் பயன்தரத்தக்க சிறந்த வழிமுறைகளைக் கடைப்பிடிப்பது இன்றியமையாததாகும். அதற்கான செலவும் அதிகரிக்கும்.

குறிப்பாக, இறுதிகட்ட முடிவுகளை அடையும் நிலையில் மாதிரிமுறை சாராத பிழைகளை அறவே நீக்குவதற்காக, அதிக அளவு முயற்சியும், (சிரமமும்) செலவும் கருதப்படவேண்டி யிருக்கும். பெரிய அளவில் ஏற்படும் (major) தவறுகளை அல்லது பிழைகளைக் கண்டறிந்து அவற்றைக் கட்டுப்படுத்துவதானது, மாதிரி முறைச் சரிபார்த்தல்மூலம் எளிதாக அமைய வாய்ப்பு உண்டு. ஆனால் மிகச் சிறிய அளவில் (minor) ஏற்படும் பிழைகளின் மூலக்கூறுகளை அறிந்து திருத்தங்கள் செய்வது அவ்வளவு எளிதல்ல, அதற்கு ஆகும் செலவும் அதிகம். எனவே, தேவைப்பட்ட நோக்கத்திற்கு ஏற்றபடி முடிவுகளைப் பெறத்தக்க அளவில், முடிந்தவரை இத்தகைய பிழைகளைக் குறைக்க முயலுவதே அறிவுடைமையாகும். பின்வரும் சில வழிமுறைகள் பெரும்பாலும் கையாளப்படுவன ஆகும் :

(i) விவரங்களின் பொருத்தத்தைச் சரிபார்த்தல் (Check on Consistency)

சேகரிக்கப்படும் விவரங்களின் தன்மையை அறியும் பொருட்டு, வினாப்பட்டியலினை உருவாக்கும் நிலையிலேயே சில விவரக் கூறுகளை (items of information) சேர்த்துக்கொள்வது சாலச் சிறந்தது. இவ்விவரக் கூறுகளைப் பெறுவது எளிதாயின், அவற்றை ஆய்வின் மற்ற உறுப்புகட்கும் (units) விரிவுபடுத்த லாம். உதாரணமாக, மக்கட்தொகைக் கணிப்பு ஆய்வுகளில் (census surveys on population) பயன்படுத்தப்படும் இரு முறைகளைக் கருதுவோம். அவை (i) 'டி ஜூரி' முறை (de jure), (ii) டி ஃபாக்டோ முறை (de facto) என்பனவாகும்.

முதல் முறையில், கணிப்பு நடைபெறும் இரவில் ஒரு நபர் தங்கியிருக்கும் இடத்தில் கருதப்பட்டுக் கணக்கெடுக்கப்படுவார். மாறாக, அவர் கணிப்பின் ஒரு குறிப்பீட்டு காலத்தில் (reference period) இயல்பாக வசிக்கும் இடத்தில் (living place) கணக் கெடுக்கப்பட்டால், அது இரண்டாவது முறையாகும். முதல் முறை மேற்கொள்ளப்படின், அங்கு இரண்டாவது முறையினையும் கருதுவதன் மூலமாக, எவ்வளவு பேர் தற்காலிகமாக இருக்

கின்றனர், எவ்வளவு பேர் தற்காலிகமாக இல்லாதிருக்கின்றனர் (absent) என்பன பற்றிய விவரங்கள், பெறப்படும் தகவல்களின் தன்மையினை நன்கு அறிய மிகவும் உதவிபுரிகின்றன. சில நிலையான விகிதங்கள் பற்றிய விவரங்களும் இதற்கு மிக்க துணை புரிகின்றன.

மற்றொரு முறையானது, சந்தேகத்திற்குரிய விவரங்களை அறிய உதவுகிறது. உதாரணமாக, குடும்பங்களின் செலவு பற்றிய ஆய்வினில், எஃகுதலா மொத்தச் செலவானது (per capita expenditure) கணக்கிடப்பட்டு, ஒவ்வொரு மாதிரிக் கூறினுக்கும் தக்கபடி ஒரு வரைபடத்தில் காட்டப்படுவதின்மூலம், எதிர்பார்க்கப்படும் ஒரு போக்கினின்று (pattern) விலகி அமையும் ஏதேனும் ஒரு நிலையில், அத்தகைய மதிப்புகளை மீண்டும் கண்டறிந்து சரிபார்க்க வழி ஏற்படுகிறது. ஆய்வாளரின் மூலம் பெறப்படும் சில விவரங்களின் நம்பகத் தன்மையும் சரிவர சோதித்து அறியப்படுகிறது.

(ii) மாதிரிச் சரிபார்த்தல் முறை (Sample Check)

ஏதேனும் ஒரு ஆய்வின் பல்வேறு நிலைகளிலும், மீண்டும் ஒருமுறை ஆய்வு நடத்துவதன்மூலம் பல பிழைகளைக் கண்டறியவும் அவற்றை நீக்கவும் இயலும். ஆனால் செயல் முறைக்கு இது எளிதான ஒன்றல்ல. எனவே, முழு அளவில் ‘மறு ஆய்வு’ நடத்தப்பெறுமல், நடத்தப்பட்ட ஆய்வின் ஒரு மாதிரியினைக் கருதி, தெரிந்தெடுக்கப்பட்ட சில ஆய்வாளரைக் கொண்டு விவரங்களின் ஏற்புத் தன்மை சரி பார்க்கப்படுகிறது. மாதிரியினைத் தக்கவழியில் தேர்ந்தெடுத்து, மிகத் திறமையுடன் சரிபார்க்கும் முறையானது அமைந்திடப் பெறின், மாதிரிமுறை சாராத பிழைகளின் தெளிவான மதிப்பீடு பெறப்படும்.

இதன் பொருட்டு, வெவ்வேறு நிலைகளில், முடிவுகளைத் தகுந்த முறையில் பாகுபாடு செய்து, தொகுப்புகளை கருதி (lots) ஒவ்வொரு தொகுப்பினின்றும் மாதிரிகளைத் தேர்ந்தெடுத்து ‘மறு ஆய்வு’ நடத்தலாம். அதாவது, பூர்த்தி செய்யப்பட்ட பட்டியல்கள், குறியிடப்பட்ட பட்டியல்கள், கணக்கீட்டு ஏடுகள் போன்ற பலவற்றை தொகுப்புகளாகக் கொண்டு, ஒவ்வொன்றி லிருந்தும் மாதிரித் தேர்வு (sample selection) நடத்தப்படுகின்றது. நிலைமைக்கு ஏற்றவாறு இம்முறையானது பயன்படுத்தப் படவேண்டும்.

(iii) ஆய்வு முற்றுப்பெற்ற நிலையில் நடத்தப்படும் சரிபார்த்தல் (Post-Survey Check)

ஒரு மாதிரி ஆய்வினில் (அல்லது முழுக்கணிப்பில்), ஆய்வு நடத்தப்பட்டு முடிக்கப்பட்ட கூறுகளினின்று ஒரு துணைக் கூறினை (மாதிரியினை) (subsample) (அல்லது மாதிரியை) தேர்ந்தெடுத்து மறு ஆய்வு அல்லது மறு பரிசீலனை செய்வது, மாதிரி முறை சாராத பிழைகளைக் கணக்கிட்டு மதிப்பிட நடைமுறையில் மேற்கொள்ளப்படும் முறை ஆகும். இதற்காக, நன்கு பயிற்சி பெற்ற, அனுபவம் வாய்ந்த பணியாளர்கள் (staff) பயன்படுத்தப் படுகின்றனர். இத்தகைய சரிபார்ப்பு—ஆய்வு அல்லது ‘ பரிசீலனை—ஆய்வு ’ (check-survey) சரிவர நடைபெறுதற்கு, ஆய்வு நடத்தி முடிக்கப்பட்ட உடனே இத்தகைய மறு பரிசீலனையைக் கருதுவதும், மூல ஆய்வின் கட்டுப்படுத்தும் விளைவினைக் குறைக்கதக்க வழிமுறைகளைக் கருதுவதும் அவசியமாகும்.

பரிசீலனை ஆய்வுகளைத் தக்கமுறையில் திட்டமிடல் வேண்டும். அதன்மூலம், ஆய்வு நடத்தப்பட்ட பரப்பு, மற்றும் பெற்ற விவரங்கள் இவற்றின் பிழைகளை நன்கு மதிப்பிட இயலுதல் வேண்டும். இதன்பொருட்டு, முதலில், பெரிய அளவில் கருதப்படும் பரப்புக் கூறுகளினின்று எடுக்கப்படும் மாதிரியில் உள்ள உறுப்புகளை நன்கு ஆராய்தல் வேண்டும். அதாவது, கருதப்படும் கிராமம், கணிப்பு நடைபெறும் பரப்பு போன்றவை பற்றியதாகும். இதன்பின், இறுதிநிலைக் கூறுகளினின்று ஒரு மாதிரியினை எடுத்து, எல்லாப் பகுதிகட்கும் பிரதிநிதித்துவம் அளித்து, அதன்மூலம் மறு ஆய்வானது நடத்தப்படும்.

இத்தகைய மறு ஆய்வானது சிறந்த திறன்பெற்று விளங்க வேண்டுமெனின், முதல்முறையாகக் கருதப்பட்ட மாதிரிக் கூறுகளின் பட்டியலும், விரிவான, பூர்த்தி செய்யப்பட்ட வினாப் பட்டியல்களும், பரிசீலனைக்கு எடுக்கப்பட்ட கூறுகளின் பட்டியலும், ஆய்வாளரிடம் கொடுக்கப்பட்டு, முதலில் கண்டறிந்த விவரங்களுடன் அவரது முடிவுகள் ஒப்பிடப்பட, பிழைகளை மதிப்பிடுவதும் எளிதாக அமையும். மேலும், முதலில் இல்லாத சில பிழைகளைச் செயற்கையாக, பட்டியலினிடையே சேர்ப்பதன்மூலம், ஆய்வாளரை மிக்க கவனத்துடன் பணியாற்றும்படிச் செய்யலாம்.

இத்தகைய சரிபார்ப்பு ஆய்வு மூலமாக, இருவேறு விதங்களில் கணிப்பு மேற்கொள்ளப்பட்ட கூறுகளைப் பொறுத்த வரையில், அவற்றின் விவரங்களைத் தனித்தனியாக ஒப்பிட்டுச் சரிபார்த்து, ஏற்படும் வேறுபாடுகளை ஆராய முடிகிறது. வேறுபாடுகள் காணப்படின், அவற்றின் காரணத்தினை நன்கு கண்டறிய முயற்சி மேற்கொள்ளப்பட்டு, அதன் விளைவாக மாதிரி சாராத பிழைகளின் தன்மைபற்றிய ஒரு மதிப்பீடும் கிடைக்கின்றது. பெரிய அளவில் விவரங்களை இணைத்து ஒப்பிடுவது நடைமுறைச் சிக்கல் நிறைந்த ஒன்ருகும் என்பதும் கவனிக்கத்தக்கது. இரண்டு விவரத் தொகுதிகளை இணைப்பதில் காலவிரயமும், பணவிரயமும் ஏற்படுகின்றன.

தனித்தனியாக விவரங்களை ஒப்பிடுவதைவிட, மொத்த அடிப்படையில் (aggrgate) சரிபார்க்கும் முறையானது கையாளுதற்கு எளிதாக அமைகிறது. இதன்மூலம், பரிசீலனை ஆய்வு விவரங்கள்மூலம் பெறப்பட்ட மதிப்பீடுகள், முதன்முதல் மேற்கொள்ளப்பட்ட ஆய்வில் பெற்ற மதிப்பீடுகளுடன் ஒப்பிட்டுச் சரிபார்க்கப்படுகின்றன. இதன் விளைவாகப் பெறப் படும் பிழை மதிப்பீடானது, மாதிரிமுறை சாராத பிறழ்ச்சியினை (non-sampling bias) பற்றியதாக அமையும். ஆனல், தனித்தனி யாக விவரங்கள் ஒப்பீடு செய்யப்படுகையில், அதன்மூலம் மதிப்பிடப்படும் பிழை அளவு, மாதிரி சாராத மாறுபாட்டினைப் பற்றி அமையும்.

மேலும், பரிசீலனை ஆய்வின் பயன்படுத்தப்படும் மாதிரியின் அளவானது (sample size), ஒரு குறிப்பிட்ட நம்பகத்தன்மையில், ஏதேனும் ஒரு குறிப்பிட்ட மதிப்பைவிட அதிகமாக அமையும் பிழையினைக் கண்டறியத்தக்க விதத்தில் அமைதல்வேண்டும். மாதிரியின் வடிவ அமைப்பு (design), பின்பற்றப்படும் சரிபார்ப்பு முறை, மதிப்பீட்டுத் திட்பம் ஆகிய பலவற்றைப் பொறுத்து மாதிரி அளவானது அமைகிறது. தகுந்த அளவில் ஒரு முன்னேடிக் கணிப்பாய்வு (pilot survey) நடத்துவதன்மூலமாக, மாதிரி அளவினைத் தீர்மானிக்க உதவும் விவரங்களைப் பெற வழி ஏற்படும். இத்தகைய பரிசீலனை ஆய்வின் முடிவுகளையும் இறுதியில் வெளியிடுதல் அவசியம்.

(iv) வெளிக் குறிப்பேடுகளைக் கொண்டு சரிபார்த்தல் : (Check with External Records)

வேறு ஏதேனும் மூலக்கூறுகளினின்று மாதிரியினைத் தேர்ந் தெடுத்து, கருதப்பட்ட ஆய்வினில் எல்லாக் கூறுகளும்

கணிக்கப்படுகின்றனவா எனவும், ஏதேனும் வேறுபாடுகள் (discrepancies) காணப்படுகின்றனவா எனவும் கண்டறியலாம். இதற்காகப் பயன்படுத்தப்படும் விவரம் முழுமையாக இருக்க வேண்டுமென்ற கட்டாயமில்லை. உதாரணமாகக் மக்கட்தொகை கணிப்பாய்வுகளில், பல்வேறு மதிப்புகளைச் சரிபார்த்து அறிவதற்காக, மறு ஆய்வில் (பரிசீலனை ஆய்வு), பிறப்புகள்பற்றிய பதிவேடுகள், பள்ளி செல்லும் சிறுவர் சிறுமியர் எண்ணிக்கை, வயதானோர் எண்ணிக்கை போன்ற பல்வேறு வகைப்பட்ட வெளி விவரங்கள் பயன்படுகின்றன. சில கூறுகளை நன்கு தெரிந்து, அவற்றின் இடமறிந்து ஆய்வு நடத்துவது எளிதாக அமைவதில்லை; எனவே இம்முறையானது செயல்முறைக் கடினமுடையதாகும்.

(v) தரக் கட்டுப்பாட்டு முறை : (Quality Control technique)

பெரிய அளவில், மீண்டும் மீண்டும் வருவனவாக அமைந்த செயல்முறைகள் ஆய்வில் இடம்பெறுவதால், புள்ளி விவர தரக் கட்டுப்பாட்டு முறைகள் அதிக அளவு பயன்பெறுகின்றன. கட்டுப்பாட்டு விளக்கப் படங்கள் (control charts) மற்றும் ஒத்துக்கொள்ளும் மாதிரித் தேர்வு வழிமுறைகள் (acceptance sampling techniqes) இவற்றை, விவரங்களின் தரத்தை மதிப்பிடும் பொருட்டு மிகவும் பயன்தரத்தக்க அளவில் மேற்கொள்ளலாம். பல்வேறு முறைகளில், சரியான நடவடிக்கையினை மேற்கொள்ளத் தகுந்த விதிகளை அமைத்துக் காட்டும் வழிமுறையானது மிகவும் கருதப்படும். அதன்படி, ஏதேனும் ஒரு ஆய்வினில், ஒவ்வொரு ஆய்வாளரின் பணியினையும் தக்கமுறையில் சரிபார்த்து, பிழை வீதமானது மதிப்பிடப்படும். பிழை ஏற்படும் வீதம் (error-rate) குறைந்திடின், ஆய்வாளரின் பணியினில் ஒரு மாதிரித் தேர்வின் அடிப்படையில், சரிபார்த்தல் நடைபெறும். இத்தகைய வழிமுறைகள் மூலமாக, சரிபார்த்தலில் ஏற்படும் செலவு குறைவதுடன், இறுதி முடிவுகளின் தரமும் பெரிதும் உயர்ந்திருக்கும்.

11. அளவீட்டுப் பிழைகள்-ஒரு குறிப்பு (Errors in Measurement)

ஒவ்வொரு மாதிரி மதிப்பும் (sample observation), அளவீட்டுப் பிழைகளினால் பாதிக்கப்பட வாய்ப்பு உண்டு. எப்பொழுதுமே, மாதிரியில் கண்டறியப்பட்ட மதிப்பு, சரியான மதிப்பிற்குச் சமமாக இருக்கும் என்று எதிர்பார்க்க இயலாது.

பெரிய அளவினதான மாதிரிகளில், வெவ்வேறு தனிப்பிழைகள் அனைத்தும் இறுதியில் தம்முள் ஒன்றோடொன்று அடிபட்டுப் போகலாம் (cancel out) என்று கருதலாம். அதன்மூலமாக, மாதிரி மதிப்புகளின் சராசரியானது, உண்மை மதிப்புகளின் சராசரிக்கு (ஏறக்குறைய) சமமாக அமையலாம்.

கணித அடிப்படையில் ஒரு நியதியினை வரையறுப்போம் :

α-ஆவது திருப்பத்தில் (repetition) i என்ற மாதிரிக் கூறின் (sampling unit) கண்டறியப்பட்ட மதிப்பினை $u_{i\alpha}$ என்று கொள்வோம். இது, அக் கூறினது உண்மையான மதிப்பிலிருந்து (true value) மாறுபடும். உண்மை மதிப்பு u_i என்றும், பிழையினை $\in_{i\alpha}$ என்றும் குறிப்பிட்டு,

$$u_{i\alpha} = u_i + \in_{i\alpha} \qquad \ldots\ldots \quad (1)$$

என்று வரையறுக்கிறோம்.

மாதிரியானது சமவாய்ப்பின் அடிப்படையில் அமைவதாகக் கொண்டால்,

$$E_i\,(u_{i\alpha}) = u_i$$

$$E_i\,(\in_{i\alpha}) = 0$$

மற்றும் $E\,(u_{i\alpha}) = E\,(u_i) = \mu$

என்று எதிர்பார்க்கிறோம். μ என்பது முழுமைத்தொகுதி சராசரியாகக் கருதப்படும். E_i என்பது, i-ஆனது கொடுக்கப்பட்டு, எடுக்கப்படும் கட்டுப்பாட்டு எதிர்பார்க்கும் மதிப்பாகும் (conditional expectation). E என்பது கட்டுப்படுத்தப்படாத (unconditional) எதிர்பார்க்கும் மதிப்பாகும்.

மேற் குறிப்பிட்ட வரையறை (1)-ல், $\in$ என்பது ' மாதிரி முறை சார்ந்த ' பிழையாகும்.

மாதிரிமுறை முற்றிலும் சமவாய்ப்பின் அடிப்படையில் அமையாத நிலையில், மற்றொரு வகைப் பிழையான ' பிறழ்ச்சி ' (bias) ஏற்படுகிறது. இத்தகைய பிழை (மாதிரிமுறை சாராமல்),

வாய்ப்பினோடு தொடர்பின்றி (non-stochastic), அதே சமயம் ஒவ்வொரு மாதிரிக் கூறினையும் பாதிக்கும் விதமாக அமைகின்றது. இந் நிலையில், வரையறையானது

$$u_{i\alpha} = u_i + \beta_i + \in_{i\alpha} \qquad \ldots\ldots \quad (2)$$

என்று அமையும்.

இதில் $\in_{i\alpha}$ என்பது, வாய்ப்பினோடு (chance) தொடர்புற்று அமையும் பிழையாக, $E_i(\in_{i\alpha}) = 0$ என்றவாறு காணப்படும். மேலும் ' β_i ' என்பது $E(\beta_i) = \beta$ என்று அமைந்திருக்கும் மாறுத (constant) ' பிறழ்ச்சிக் கூறுகும் (bias component). ஆகும்.

எனவே, $E_i(u_{i\alpha}) = u_i + \beta_i$

மற்றும் $E(u_{i\alpha}) = \mu + \beta$

என்று முடிவுகள் அமைகின்றன.

இதன்மூலம், $[E(u_{i\alpha}) - \mu] = \beta$ என்பது பொதுவாக 'பிறழ்ச்சி' என்று கூறப்படுகிறது. இது நேர் கணியமாகவோ(+) அல்லது எதிர்க்கணியமாகவோ (—) இருக்கலாம். இது எப்போதும் மாறுத பிழைக்கூறுக அமைந்திடும். மாதிரி அளவு அதிகரிக்கையில், இப் பிழையானது குறையாது.

இத்தகைய மாறுப்பிழையானது (bias), மதிப்பீடுகளின் மாறுபாட்டு அளவினைப் பொதுவாக பாதிப்பது இல்லை என்பது குறிப்பிடத்தக்கது.

' n ' அளவுள்ள மாதிரியின் சராசரியின் மாறுபாடு

$$V(\bar{u}_n) = V(\bar{u}) + V(\bar{\beta}_n) + V(\bar{\in}_n)$$

$$\therefore \quad V(\bar{u}_n) = \frac{\sigma_u^2}{n} + \frac{\sigma_{\in}^2}{n}$$ என்றுகும்.

($\in$ மற்றும் u இவை தொடர்பற்றவை என கருதப்பட்டது).

12 பதில் அளிப்பில் பிழைகள் பற்றி ஒரு குறிப்பு (Errors in Response)

நேரடித் தொடர்பின் தன்மை, ஆய்வாளரின் அறிவுத்திறன், பதிலளிப்போரது நிலை ஆகிய பல்வேறு காரணங்களால் பதில் அளிப்பு பிழைகள் (Response errors) ஏற்படலாம். இவற்றோடல்லாமல், சில விவரங்களை மீண்டும் நினைவுகூர்கையில் சில பிழைகள் ஏற்பட நேரிடும். இவை பதிலளிப்போரின் கட்டுப்பாட்டினையும் மீறும் தன்மையன ஆதலால், இவற்றின் மூலம் ஏற்படும் விளைவுகளை சற்று கருதுவது அவசியம். இத் தகைய பிழைகளை 'மீண்டும் கவனத்துக்குக் கொண்டுவருகையில் நிகழும்' பிழைகளாகக் (Recall Error) குறிப்பிடலாம். ஆய்வின் அறிவிப்புக் காலத்திற்கும், கணிப்பு நடத்தப் பெறும் காலத்திற்கும் இடையில் உள்ள இடைவெளியின் அளவினைப் பொறுத்து இத்தகைய பிழைகளின் தன்மையானது அமைந்திருக்கும்.

இதனை ஆராய்தற்கு, ஆய்வில் கருதப்படும் துணைமாதிரி (அல்லது மாதிரி) யின் கூறுகளைப் பல்வேறு அறிவிப்புக் காலங்களுடன் தொடர்பு படுத்தி ஆய்வதன் மூலமோ, அல்லது சில துணை விவரங்களைப் பெறுவதன்மூலமோ, செயல்முறை எளிதாக்கப் படுகிறது,

ஒரு அறிவிப்புக் காலவரையில் பெறப்பட்ட விவரங்கள், மற்ற காலவரைகளில் அறிவிக்கப்படும் விவரங்களை கட்டுப் படுத்துவதானது, பிரச்சினைகளை உருவாக்குகிறது. வெவ்வேறு அறிவிப்புக் காலங்களில் பெற்ற விவரங்கள் ஒன்றோடொன்று ஒத்த தன்மையில் அமையாவிடின், எந்தக் காலத்தைக் கருத வேண்டும் என்று தீர்மானிப்பது சிக்கலான பிரச்னையாகும். இதற்கு தக்கபடி ஒரு கட்டுப்பாட்டு மாதிரியினைக் கருதி, நம்பகத் தன்மையுடைய விவரங்களைப் பெறும்பொருட்டுத் தனி முயற்சி எடுக்கப்படலாம். அல்லது, கருதப்படும் வெவ்வேறு காலங்களுக் கேற்ற வகையில் விவரங்களை, வெவ்வேறு மாதிரிக் கூறுகளைக் கருதுவதன்மூலம் சேகரிக்கலாம். இதற்கு, எடுக்கப்படும் மாதிரியின் அளவானது பெரிதாக அமையவேண்டும். இதனைத் தவிர்க்க, ஒன்றோடொன்று தொடர்புபடுத்தப்பட்ட துணை மாதிரிகளைக் கருதி வெவ்வேறு அறிவிப்புக் காலங்களில் விவரங்களைப் பெறலாம்.

அல்லது, மற்றொரு முறையில், சில துணை விவரங்களைச் (additional information) சேகரித்து, அவற்றின்மூலம்

வெவ்வேறு அறிவிப்புக் காலங்களுக்கு ஏற்ற மதிப்பீடுகளைப் பெற்று, அவற்றைப் பயன்படுத்துவதன் மூலமாக, மீண்டும் விவரங்களை நினைவுகூர்தலில் ஏற்படும் தவறுகளை (பிழைகளை) ஆராயலாம். பிறப்பு பற்றிய கணிப்பு ஆய்வுகளில், ஒரு இடத்தில் நிகழும்பிறப்பு வீதத்தினை கணக்கிட ஒரு குறிப்பிட்ட ஆண்டினில் நிகழும் பிறப்பு எண்ணிக்கையினை மட்டுமன்றி, எடுக்கப்படும் மாதிரிக் கூறுகளைப் பொறுத்தவரை, ஒரு குறிப்பிட்ட மாதம் அல்லது நாளில் நிகழும் பிறப்புகளின் எண்ணிக்கையை அறிவதன் மூலமாக, நிகழும் காலத்திற்கும் அறிவிக்கும் காலத்திற்கும் இடையில் உள்ள இடைவெளிக்கேற்ப பிறப்பு வீதத்தினைப் பாகுபாடு செய்து இத்தகைய பிழைகளை அறிவது எளிதாகும்.

முடிவுரை :

மாதிரிமுறை சாராத பிழைகளைப் பற்றியும் அவற்றின் விரிவான ஆய்வு முறைகளையும் இதுவரையில் விளக்கமாகக் கண்டோம். பொதுவாக, அறியப்படும் விவரங்களின் அல்லது பெறப்படும் முடிவுகளின் தரத்தினை (quality) கட்டுப்படுத்துவதே அடிப்படை நோக்கம் என்று காணலாம். தொடர்பு கொள் வோரிடை (interviewer) பிறழ்ச்சி வேறுபாடுகள், விவரங்களைத் தெரிந்து கொள்வதில் கையாளப்படும் வெவ்வேறு வழிமுறைகள், போன்றவற்றைப்பற்றி அறிந்து, தவறுகளைப் பற்றிய ஒரு மதிப்பீட்டினைப் பெறுதற்கு, மஹலநோபிஸ் (Mahalanobis) கையாண்ட ஒரு வழியினைப் பற்றிய குறிப்புடன் இப்பகுதியினை முடித்திடுவோம்.

அதாவது ஒரே மாதிரியினை, இரண்டு அல்லது அதற்கு மேற்பட்ட துணைக்கூறுகளாக அமையும்படி தெரிந்தெடுக்கும், 'ஒன்றோடொன்று ஊடுருவும் தன்மையுள்ள' துணை மாதிரிகள் எடுக்கப்படும் முறையாகும். எல்லாத் துணை மாதிரிகளும் ஒரே அடிப்படையில் தேர்ந்தெடுக்கப்படுகின்றன. இந்த துணை மாதிரி களில், வெவ்வேறு ஆய்வாளர்கள் மூலம் ஆய்வு நடத்தப் பெற்று, வெவ்வேறு குழுவினரால் செயல்முறைப்படுத்தப்பட்டு, இறுதி முடிவுகள் ஒன்றோடொன்று ஒப்பிடப்படுகின்றன. ஒவ்வொரு நிலையிலும் துணைக்கூறுகளே கருதப்படும்.

விவரங்களின் தரத்தினை மதிப்பிட, ஓரினப் படுத்தப்பட்ட கூறுகளைக் கருதி, திரள்களை அமைத்து அவற்றினின்று துணைக் கூறுகள் எடுக்கப்படவேண்டும். இதன்மூலம், துணைமாதிரி மதிப்பீடுகள் அதிகத் திறனுடன் ஒப்பிடப்படுகின்றன. மக்கட்

தொகைக் கணிப்பு ஆய்வுகளில், ஒன்றுவிட்டு ஒன்ருக அமைந்த கூறுகளை ஓர் ஆய்வாளர் கருதுவதாகவும், மீதி உள்ளவற்றை மற்ருெருவர் கருதுவதாகவும் கொண்டு ஆய்வு நடத்தப்படின், அவர்களிடையே உள்ள பிறழ்ச்சிகளின் வேறுபாடு அறியப்படும்.

ஆய்வாளரிடை அமையும் மாறுபாடு, அட்டவணைப்படுத்து வோரிடை உள்ள மாறுபாடு, ஒன்ருேடொன்று ஊடுருவும் நிலையில் எடுக்கப்பட்ட துணைமாதிரிகள்மூலம் காணப்படும் உடன் விளைவுகள் போன்ற பலவகையான பிழைக்கூறுகளைத் தக்கபடி பிரித்துப் பகுப்பாய்வு நடத்தலாம். மாறுபாட்டுக்கூறு களின் பகுப்பாய்வு முறைகள் (analysis of variance) இதற்கு பெரும்பாலும் உதவுகின்றன. அட்டவணைப்படுத்தும் குழுக்களின் இடைப்பட்ட மாறுபாடு, ஆய்வாளர் குழுக்களிடை அமையும் மாறுபாடு, மீதி உள்ள மாறுபாடு போன்றவற்றைத் திறனுடன் ஆராய்ந்திட இயலுகிறது. இதன்மூலம் மாதிரி சாராத பிழைகளைத் தக்கபடி அறிந்து, பல்வேறு நிலைகளிலும் அவற்றின் விளைவைக் குறைக்க வழி ஏற்படும்.

11. இந்திய நாட்டுப் புள்ளிவிவரங்கள்
(Indian Statistics)

[துவக்க நிலை—நாட்டு விடுதலைக்குப்பின்—சில அமைப்புகள்—புள்ளி விவரங்களின் சில வகைகள்—தேசிய மாதிரி அளவெடுப்பு—மக்கட் கணிப்பு ஆய்வு—விவசாயப் புள்ளி விவர ஆய்வு—தொழிற் புள்ளி விவர ஆய்வுகள்—வர்த்தகப் புள்ளி விவர ஆய்வு.]

1. துவக்கநிலை :

இந்திய நாட்டின் பொருளாதார, சமூக, வாணிப முன்னேற்றத்திற்கும், நாட்டின் வளர்ச்சியினைக் குறிக்கோளாகக் கொண்டு திட்டங்கள் மேற்கொள்ளுதற்கும், பல்வேறு துறைகள் பற்றிய புள்ளிவிவரங்கள் சேகரிக்கப்பட்டு, விரிவான ஆய்வுகள் நடத்தவேண்டியது இன்றியமையாததாகும். பன்னெடுங்காலமாக இவை கருதப்பட்டு வருவதை வரலாறு கூறும். முன்னுளில், நாட்டை ஆட்சி புரிவோர்க்கு. நாட்டு மக்கட்தொகை, நிதி நிலைமை,, விவசாய மற்றும் தொழில் உற்பத்தி, படைகளின் நிலை போன்ற பல்வேறு இயல்புகள் பற்றிய புள்ளிவிவரங்கள் பெரும் அளவில் தேவைப்பட்டன. சமுதாய வளர்ச்சிக்குத் திட்டங்கள் உருவாக்குவதும், எதிரிகளைச் சமாளிக்க நடவடிக்கை எடுப்பதும் போன்ற பலவற்றிலும் இவை பேருதவி புரிந்தன. வெள்ளையர் ஆதிக்கத்தில் (British regime), நம் நாட்டில் புள்ளிவிவர ஆய்வுக்குத் தனியான கவனம் செலுத்தபடவில்லை. அன்றுட அலுவல்களில் ஒரு பகுதியாகவே, சில அமைப்புகள் (organisations) அரசினரால் ஏற்படுத்தப்பட்டன.

நாளடைவில். புள்ளிவிவரங்களின் அவசியம் உணரப்பட்டு, சில குழுக்கள் (Committees) உருவாக்கப்பட்டன. (Economic Enquiry Committee; Bowley-Robertson Committee; Inter Departmental Committee போன்றவை). புள்ளி விவரங்களை சேகரித்து ஆய்வுகள் மேற்கொள்ளும் பொருட்டு, நிரந்தர அமைப்பாக, 'இந்திய அரசின் பொருளாதார ஆலோசகர் அலுவலகம்' (Office of Economic Adiviser to Govt. of India) என்பது அமைக்கப்பட்டது (1938). பொது இயக்குநரின் தலைமையில், ஆய்வுப்பணி மேற்கொள்ளப்பட்டது.

நாட்டு விடுதலைக்குப்பின் (After Independence) :

விடுதலை பெற்று முழுஉரிமையுடன் செயலாற்றத் துவங்கிய பின்னர், பல்வேறு ஆக்கப்பணிகட்கும் பொருளாதாரத் திட்டங்கள் அமைக்க முற்பட்ட நிலையில், இந்திய அரசானது, முழுஅளவில் தேவையான புள்ளிவிவரங்கள் சேகரிக்கப் படவேண்டியதன் அவசியத்தை நன்குணர்ந்து, சமூச வாழ்வின் ஒவ்வொரு இயல்பு குறித்தும் விவரங்களைப் பெறுவதற்குப்போதிய நடவடிக்கைகளை மேற்கொண்டது. விரிவான அளவில் பல்வேறு துறைகளிலும் புள்ளிவிவரங்கள் திரட்டப்படுவதைத் தற்கால சூழ்நிலையில் காணலாம். மத்திய, மற்றும் மாநில மட்டங் களில் ஆய்வுகள் மேற்கொள்ளப்பட்டு, விவரங்கள் சேகரிக்கப் படுகின்றன. பல்வேறு வகைப்பட்ட ஆய்வுகளின் முடிவுகளும் ஒருங்கு இணைக்கப்பட்டு (Coordinated), பலவிதமாகப் பயன் படுத்தப்படுகின்றன. மாநிலங்களில் நடத்தப்படும் (விவசாயம், கால்நடை, பொதுநலம், வனவளம் போன்ற சிறு அளவிலமைந்த துறைகளில்) புள்ளிவிவர ஆய்வுகட்குச் செயல்முறைகளையும், ஆலோசனைகளையும் மத்திய அரசானது வழங்குகிறது. பெரும் அளவில் மேற்கொள்ளப்படும் (மக்கட்கணிப்பு, தேசீயப் பாதுகாப்பு, தகவல்தொடர்பு, ரயில்வே போன்ற பலவகையான) புள்ளிவிவர ஆய்வுகள் மத்திய அரசினரால் (Central Govt,) நேரடியாகவே மேற்கொள்ளப் படுகின்றன. எனினும், இருவகை அரசினரும் இணைந்து செயலாற்றுவது மிகவும் இன்றியமையாத தாகும். நாடு முழுதும் கருதப்பட்ட அளவில், இறுதி முடிவுகளை மத்திய அரசானது வெளியிடுகிறது.

இவ்வாறு மாநில மற்றும் மத்திய மட்டங்களில் (State & Central levels) மேற்கொள்ளப்படும் ஆய்வுகளின் முடிவுகளை இணைக்கும் பொருட்டு, மத்திய அமைச்சு செயலகத்தின் (Cabinet Secretariat) ஒரு பகுதியாக 'புள்ளிவிவரத் துறை' (Dept. of Statistics) ஏற்படுத்தப்பட்டது. இதன் செயல்முறை (technical) பிரிவாக 'மத்திய புள்ளிவிவர ஆய்வுக் கழகம்' (Central Statistical Organisation — C. S. O) விளங்குகிறது. மேலும், வெவ்வேறு மாநிலங்களிலும் (States) புள்ளியியல் ஆய்வுக் கழகங்கள் (Bureaus) அமைக்கப்பட்டு, மாநில அளவில் நடைபெறும் ஆய்வுப் பணிகட்குப் பொறுப்பேற்கின்றன.

சில அமைப்புகள் (Organisations) :

ஒவ்வொரு அமைச்சகமும் (Ministry) ஒரு புள்ளிவிவர ஆய்வுப் பிரிவினைக்கொண்டு விளங்குகிறது. அன்றாட அலுவலின்

ஒரு பகுதியாக புள்ளிவிவரங்கள் ஒவ்வொரு துறையிலும் [வருவாய், ரயில்வேக்கள், தகவல் தொடர்பு, மொத்த அளிப்பு (supply) போன்று] தேவையான விவரங்கள் பெறப்படுகின்றன. வேறுசில அமைப்புகளும், உற்பத்தி மற்றும் பங்கீடு இவற்றோடு தொடர்புபெற்று, விவரங்களைத் திரட்டுகின்றன.

இவற்றோடல்லாமல், சில செயல் அமைப்புகள் அரசினரால் ஏற்படுத்தப்பட்டு, புள்ளிவிவரங்களைத் திரட்டி, ஆய்வுகள் மேற்கொண்டு. முடிவுகளை வெளியிடும் பிரத்தியேகப் பணியில் ஈடுபடுகின்றன. அவையாவன:

தேசீய மக்கட் கணிப்பாய்வுக் கழகம் (National Census Organization) :

பொதுப் பதிவு அதிகாரி (Registrar General) யின் கீழ்ப் பணியாற்றும் இவ்வலுவலகமானது, 1948-ல் நிரந்தர அமைப்பாக உருப்பெற்றது. நாட்டு மக்கட்தொகைக் கணிப்பானது ஒழுங்காகவும், கவனத்துடனும் மேற்கொள்ளப்படுகிறது. 'Vital Status of India' மற்றும் 'Indian Population Bulletin' ஆகியவை இதன் வெளியீடுகளாகும்.

வணிகப்புலனுய்வு மற்றும் புள்ளிவிவர ஆய்வுப் பொதுச் செயலகம் (Directorate General of Commercial Intelligence and Statistics—D. G. C. I. S)

இவ்வலுவலகமானது, தற்போது, நாட்டின் வர்த்தகம் பற்றிய (பல்வேறு வழிகளிலும்—இருப்புப்பாதை வழி, கடல்வழி, நதிவழி என்றவாறுகக் கருதி) விவரங்களைச் சேகரித்து, தொகுத்து வெளியிடும் பொறுப்பை மேற்கொண்டதாக அமைகிறது. உள்நாட்டு, மற்றும் வெளிநாட்டு வாணிபம் பற்றிய (trade) பல வெளியீடுகள் இதன்மூலம் வெளிவருகின்றன. (*Ref :* வர்த்தகப் புள்ளி விவர ஆய்வு).

உழைப்பு பற்றிய ஆய்வுக் கழகம் (Labour Bureau) :

விலைவாசி ஏற்றம், வாழ்க்கைச் செலவு அதிகரிப்பு இவற்றின் மூலம் ஏற்படும் விளைவுகளால், உழைப்பாளரிடை (labourers) அமைதியற்ற சூழ்நிலை நிலவுவதை ஆராய்வதற்கு, 'குடும்ப வரவு—செலவுத் திட்டப்' (family budget) பகுப்பாய்வுகள் மேற்கொள்ளப்பட்டு, முடிவுகள் எடுக்கப்படுகின்றன. 'Indian Labour Statistics' (Annual), Indian Labour Journal (Monthly), Trade Unions (Indian) என்பவை இதனுடைய சில வெளியீடுகளாகும்.

இராணுவப் புள்ளி விவர ஆய்வகம் (Army Statistical Organisation) :

இந்திய அரசின் பாதுகாப்பு அமைச்சகத்தின் (Ministry of Defence) ஒரு பிரிவான இது, இராணுவத் துறையில் பல்வேறு இயல்புகளை அறியும்பொருட்டு அமைக்கப்பட்ட (1947) பிரத்தியேக அலுவலகமாகும்.

தேசீய மாதிரி ஆய்வுக் கழகம் (National Sample Survey Organisation) :

பொருளாதார சமூகத் திட்டங்களுக்கு இன்றியமையாத விவரங்களைச் சேகரித்தளிக்கும் இச்செயலமைப்பு விரிவாகப் பின்பு ஆராயப்படுகிறது:

(*Ref* : *NSS* பற்றிய விரிவாய்வு).

இவற்றைத் தவிர, சில **தனியார் துறை** (Non-Government) அமைப்புகளும், புள்ளி விவரங்களைச் சேகரித்துத் தொகுத்தளிப்பதில் உதவிபுரிகின்றன. முக்கியமாக சில :—

இந்தியப் புள்ளி விவர ஆய்வுக் கழகம் (Indian Statistical Institute) :

பெரிய நிறுவனமாக விரிந்திருக்கும் இதன் நோக்கம், பல்வேறு புள்ளியியல் முடிவுகளை நடைமுறையில் பயன்படுத்தி ஆய்வுகள் மேற்கொள்வதாகும். பல பிரச்னைகள் (பொருளாதார, சமூக என்றவாறு) நன்கு ஆராயப்பட்டு, சிறந்த முடிவுகள் எடுக்கப்படுகின்றன. திட்டப் பணிக்குச் சிறந்த அளவில் உதவி புரிகிறது. இதைப்பற்றிய விரிவான கருத்து, பின்வரும் *NSS* பற்றிய கருத்துரையில் இடம்பெறும் (*Ref* : *NSS* ஆய்வு).

இந்திய விவசாய ஆராய்ச்சிக் கழகம் (Indian Council of Agricultural Research) :

1959-ல் முழு உருவம் பெற்ற நிறுவனமாகும். விவசாயப் பொருளாதாரத் துறையில் முன்மாதிரி (pioneer) ஆய்வுகள் மேற்கொள்ளப்பட்டு முடிவுகள் எடுக்கப்படுகின்றன. பயிர் சாகுபடி மதிப்பீட்டில் சமவாய்ப்பு மாதிரி முறையை அறிவுறுத்தியது இந்நிறுவன அமைப்பேயாகும்.

ரிசர்வ் வங்கியின் ஆராய்ச்சித் துறை (Research Dept. of Reserve Bank) :

புள்ளி விவர ஆய்வுத் துறை முழு அளவில் இடம் பெறுவதாகும். வங்கியின் பல்வேறு வெளியீடுகட்கும் தேவைப் படும் புள்ளி விவரங்களை திரட்டித் தொகுத்துத் தருவது இதன் பணியாகும். நாட்டில் உள்ள வங்கிகள் (banks) பணியாற்றும் சூழ்நிலை நன்கு அறியப்படும். உலக நாடுகளுடன் தொடர்பு நிலையும் அறியப்படுகிறது. பொது நிதிநிலை (Public Finance), 'பவுன்' மதிப்பில் மாறுபாடு (Bullion markets) போன்ற பலவற்றைப் பற்றி தீவிர ஆராய்ச்சிகளும் மேற்கொள்ளப் படுகின்றன.

இன்னும் பல அமைப்புகளும் சிறந்தமுறையில் பணியாற்றி வருகின்றன.

(National Council of Applied Economic Research, Institute of Labour Research ; Gokhale Institute of Politics & Research போன்ற சிலவற்றைக் குறிப்பிடலாம்.)

புள்ளி விவரங்கள் :

மத்திய, மற்றும் மாநில அரசினரால், நாட்டின் பல்வேறு துறைகளிலும் புள்ளி விவரங்கள் திரட்டப்பட்டு, விரிவான அளவில் ஆய்வுகளும் மேற்கொள்ளப்படுகின்றன. இனி வரும் பகுதிகளில், குறிப்பிட்ட **சிலவகையான** புள்ளி விவரங்களின் இயல்புகளை மட்டும் ஆராய முற்படுகிறோம். பின்வருமாறு வரிசையில் அவை அமைகின்றன :

I	தேசீய மாதிரி அளவெடுப்புகள்
II	மக்கட் கணிப்பு ஆய்வு (பிறப்பு—இறப்பு விவரங்கள் உள்பட)
III	விவசாயப் புள்ளி விவர ஆய்வு
IV	தொழிற் புள்ளி விவர ஆய்வு
V	வர்த்தகப் புள்ளி விவர ஆய்வு

I. தேசிய மாதிரி அளவெடுப்பு

(National Sample Survey)

அறிமுகம்

[*NSS*-ன் தன்மைகள்—மாதிரி அளவெடுப்பு இயக்கச் செயலகம்—மத்தியப் புள்ளி விவர ஆய்வுக் கழகம்—இந்தியப் புள்ளி விவர ஆய்வு நிறுவனம்—*NSS*-ன் மறு அமைப்பு—*NSS*-ல் மாநிலங்களின் பங்கு—*NSS*-ல் கணிப்பு முறைகள்—*NSS* அளவெடுப்புச் செயல்முறை—*NSS* அளவெடுப்பின் பொதுவான செயல் திட்டம்—அறிவிப்புக் காலம்—நடைமுறையில் மாதிரி—விவரங்களைப் பெறுவதற்குக் கையாளப்படும் வழிமுறை—நிலம் பயன்படுத்தப்படும் விதம்பற்றிய ஆய்வு—ஆய்வாளரின் பணியில சில இடையூறுகள்—*NSS* பற்றிய சில முக்கியக் குறிப்புக்கள்.]

அறிமுகம் :

இந்திய நாட்டின் பொதுப் புள்ளி விவரங்களுள், ஒரு சிறந்த. இன்றியமையாத இடத்தைப் பெறுவது 'தேசீய மாதிரி ஆய்வு' ஆகும். '*NSS*' என்று சுருக்கமாக குறிப்பிடப்பெறும் இத்தகைய ஆய்வானது, நாட்டின் பொருளாதாரம் மற்றும், சமூக நலன் இவற்றில் முன்னேற்றம் பெறத்தக்க அளவில், தேவையான பல்வேறு புள்ளி விவரங்களைச் சிறந்த முறையிலும், ஏற்கத்தக்க விதத்திலும் பெறும்பொருட்டு, அரசினரால் 1950-ஆம் ஆண்டு ஏற்படுத்தப்பட்ட ஓர் அமைப்பு (Organisation) ஆகும் இந்தக் கூட்டமைப்பானது, மத்திய நிதி அமைச்சகத்தின், பொருளாதார விவரங்கள்பற்றிய இலாகாவில் இடம்பெற்று, அரசின் பலவகைத் தேவைகளை முழு அளவில் பூர்த்தி செய்ய முயலுகிறது.

நாட்டில் நடைபெறும் பலவகையான மாதிரி ஆய்வுகளில் மிகப் பெரிய அளவில் நடைபெறுவதாகவும், விரிந்த பரப்பில் கருதப்படுவதாகவும், தொடர்ச்சியாக (continuous) நடத்தப்படுவதாகவும் அமைவது *NSS* ஆகும். ஒழுங்கான நியதிகளின் அடிப்படையில் மாதிரி முறைகளைக் கையாண்டு, தொடர்ந்து பல சுற்று (rounds) ஆய்வுகளை நடத்தி, அவற்றின்மூலம் தேவைப்படும் விவரங்களை அனைத்தையும் பெற்று, பல்வேறு துறைகளில் செயல்முறைப்படுத்த உதவுவது *NSS*-ன் முக்கியப் பணியாகும்.

நாட்டின் பல்வேறு துறைகளின் தொடர்புடைய புள்ளி விவரங்களின் தன்மையினைத் தெளிவாக, விளக்கமாக, விரிவாக, நம்பத்தக்க விதத்தில் படம்பிடித்துக் காட்டுவது *NSS*-ன் அடிப்படை நோக்கமாகும். நாட்டு முன்னேற்றத்துக்கும், பல

துறைகளில் மேற்கொள்ளப்படும் விரிவான ஆராய்ச்சிகளுக்கும் பெரும் அளவில் *NSS* ஆய்வின் முடிவுகள் உதவிபுரியக் காணலாம். 1971 வரையிலான காலத்தில் இருபத்தைந்து சுற்றுக்களை முடித்திருக்கும் *NSS*-ஆனது, தொடர்ந்து, இயங்கி வரும் ஓர் அமைப்பாகிறது. இதன் நோக்கம் ஒரு குறுகிய எல்லை வரையறைக்கு உட்பட்டதல்ல.

1. NSS-ன் தன்மைகள் (Nature of NSS)

முதலில், இத்தகைய ஆய்வானது, பல நோக்கங்களை (multi-purpose) கொண்டதாக அமைவது குறிப்பிடத் தக்கது. ஏறக்குறைய எல்லாத் துறைகளிலும் அதன் முடிவுகள் பயன் படுத்தப்படுகின்றன. ஒரு குறுகிய எல்லையில், சில பிரச்சினைகளை மட்டுமே கருதி, குறுகிய காலவரையில் நடத்தப்படுவது அல்ல என்பது கவனிக்கத் தக்கது.

மேலும் தனித்தனியாக நடத்தப்படும் ' தனி ஆய்வு'களைவிட (uni-stage), பல நோக்கங்களை உள்ளிட்ட ஒரு ஆய்வானது, நடைமுறைச் சிக்கனம் உள்ளதாக அமைகிறது. நாட்டின் பொருளாதாரம், மற்றும் சமூக வாழ்வு போன்றவை பற்றி விரிவான அளவில், பரந்த இடப்பரப்பில் பெரும் அளவில் விவரங்கள் திரட்டப்பட்டுப் பயன்படுத்தப்படுகின்றன. எனவே, பல்வேறு ஆய்வுகளில் இவை பேருதவி புரிகின்றன எனலாம். உதாரணமாக, NSS ஆய்வுகளில் பெறப்பட்ட விவரங்களைக் கொண்டு நாட்டின் மொத்த வருமான அளவு (National income) கணக்கிடப்படுவதைக் கூறலாம்.

மேலும், நாட்டின் பல்வேறு கிராம மற்றும் நகர்ப்புறங்களைக் கருதி, மக்களின் தேவைப் பூர்த்திக்கான செலவு விவரம், இதர செலவுகள் விவரம், நிறுவனங்கள் பற்றிய விவரங்கள், கால்நடை பற்றிய விவரங்கள், பயிர் சாகுபடி பற்றிய உண்மைகள் போன்ற பல்வேறு வகையானவற்றைப் பெற்று, அவற்றின் இயல்புகளை நன்கு கண்டறிய வழி வகுக்கிறது NSS ஆய்வு. பொருளாதாரத் திட்டங்கள் வகுக்க, இவை மிகவும் உதவுகின்றன.

ஆய்வுகளை ஒவ்வொரு ஆண்டும் (yearly) முறையான வழியில் இவ் வாய்வுகள் நடத்தப்பட்டு, காலத் தொடர் வரிசையில் (times series) விவரங்கள் தரப்படுவதால், எதிர்காலப் போக்குகள் பற்றி மதிப்பிட்டு அறிய இயலுகிறது. உடனடி முடிவுகளைப் பெற, சில தற்காலிக (ad hoc) ஆய்வுகளும் மேற்கொள்ளப்படுகின்றன.

முதன்முதலில் கருதப்பட்ட சுற்றில், நாட்டின் மொத்த வருமானத்தை (National Income) கணக்கிடும் பொருட்டு, விவரங்கள் திரட்டப்பட்டன. பின்பு சிறிதுகாலம் சென்று நடத்தப்பட்ட 8-ஆவது சுற்றில், மக்கள் வசம் உள்ள நிலங்கள் பற்றிய விவரங்கள் சேகரிக்கப்பட்டன. அடுத்து வேலையின்மை (unemployment) யின் வீதமானது கருதப்பட்டது. அதற்கடுத்த சுற்றுக்களில், விவசாயப் பணியாளரின் பொருளாதார நிலை, அவரது மொத்தச் செலவு, பெறும் கூலி அளவு, போன்ற பல விவரங்கள் கண்டறியப்பட்டன. படிப்படியாக ஒவ்வொரு சுற்றிலும் ஆய்வு நிலை செம்மைப் படுத்தப்படுவதைக் காணலாம். பயிர்கள் பற்றிய ஆய்வுக்கு அதிக முக்கியத்துவம் அளிக்கப் படுகிறது.

2. மாதிரி அளவெடுப்பு இயக்கச் செயலகம் (Directorate of Sample Survey)

சிறந்த முறையில் திட்டமிட்டு, ஏற்கத்தக்க வகையில் புள்ளி விவரங்களைப் பல துறைகளில் பெறுவது இப்பிரிவின் பொறுப்பாக அமைகிறது. பெரிய அளவில் ஆய்வுகளைத் திட்டமிட்டு, களப்பணியை மேற்கொள்ளுதல் இதனைச் சார்ந்த முக்கியமான பணியாகும். பின்வரும் விவரங்கள் திட்டவட்டமாக கண்டறியப் படுகின்றன:

(i) நாட்டின் சமூகப் பொருளாதார நிலைகளைப் பற்றிய விவரங்களைச் சேகரித்தல்; மற்றும் விலைவாசிகள், பல்வேறு ஊதியங்கள், நுகர்ச்சிக்கான செலவு, விவசாய புள்ளிவிவரங்கள் போன்ற பல வகையான விவரங்களையும் கண்டறிவது.

(ii) நாட்டின் திட்டத்தினை (planning) உருவாக்கவும், தேசீய வருமானத்தைக் கணக்கிடவும் பெரிதும் உதவத்தக்க புள்ளி விவரங்களைப் பெறுதல். மேலும் வெவ்வேறு அமைச்சகங்களில் எடுக்கப்படும் முடிவுகட்கு உதவத்தக்க விவரங்களையும் பெறுவது.

(iii) திட்டமிட்டு அமைக்கப்பட்ட, நாட்டின் பல்வேறு தொழில் நிறுவனங்கள் பற்றிய தெளிவான புள்ளிவிவரங்களை அறிதல்.

(iv) விவசாயத் துறையில், பல்வேறு மாநிலங்களில் மேற்கொள்ளப்படும் ஆய்வுகட்கு துணைநின்று அவற்றின் முடிவுகளைத் தக்கபடி இணைத்து, மேலும் சில முடிவுகளை அடைதல்.

இவைபோன்ற விவரங்கள் சேகரிப்புப் பணியே இத்தகைய பிரிவின் முக்கியக் கடமையாகும்.

3. மத்தியப் புள்ளிவிவர ஆய்வுக் கழகம் (CSO—Central Statistical Organisation)

நாட்டின் மத்திய, மற்றும் மாநில மட்டங்களில் பல்வேறு வகையில் புள்ளிவிவரங்கள் கண்டறியப்பட்டு வெவ்வேறு துறைகளில் பயன்படுத்தப்பட்டு விரிவாக ஆய்வு நடத்தப்படுவதால், மத்திய அமைச்சரவையின் செயலகப் பிரிவாக, இந்த ஆய்வுக் கழகம் 1951-ம் ஆண்டு நிறுவப்பட்டது. பலவகை ஆய்வுகளின் முடிவுகளை ஒருங்கிணைத்து மேலும் விரிவான அளவில் ஆய்வு நடத்துவதும், புள்ளி ஆய்வு சம்பந்தப்பட்ட நடவடிக்கைகளை வழி வகுப்பதும் இதன் முக்கியப் பணிகளாகும். இவற்றோடு, அவ்வப்போது, முறையாகப் பலவகையான பொருளாதார மற்றும் சமூகப் புள்ளி விவரங்களை வெளியிடுவதும், தகுந்த விளக்கப் படங்களை அமைப்பதும் இதன் கடமைகளாகும். சுருங்கக் கூறின், அதிகப் பொறுப்பு வாய்ந்த இக் கழகம் நாட்டின் இன்றியமையாத ஓர் அமைப்பாக விளங்குகிறதெனலாம். புள்ளி ஆய்வுப் பணிகளில் ஈடுபடவிருக்கும் அலுவலர்கட்குச் சீரிய பயிற்சியினைச் சிறந்த முறையில் அளித்தற்பொருட்டு இவ்வமைப்பானது ஆறு அல்லது ஏழு வார காலப் பயிற்சிக் கல்வியினையும் வகுக்கிறது. மேலும், புள்ளிவிவரங்களின் அதிகப்படித் தேவையினைப் பூர்த்தி செய்தற்பொருட்டு, தனியாக ஒரு புள்ளி ஆய்வுப் பிரிவானது (Department of statistics) அரசினரால் உருவாக்கப் பட்டிருக்கிறது.

இதன் முக்கியப் பணிகளாவன :

(i) அரசின் பல்வேறு துறைகளில் நடத்தப்பெறும் ஆய்வுகளை ஒருங்கிணைத்துச் செயல்படுவதன்மூலம், ஒரே பணியானது இருமுறை செய்யப்படும் பிரச்சினை தவிர்க்கப்பட்டு, ஆய்வுச் செலவும் கணிசமாகக் குறைக்கப்படுகிறது. புள்ளிவிவர ஒருங்கிணைப்பு மிக்க பயனையும் அளிக்கிறது.

(ii) பல்வேறு கருத்துக்களை வரையறுத்து நிறுவுவதாலும், தரங்களை நிர்ணயிப்பதாலும், விரிந்த அளவில் ஆய்வுகளையும் முடிவுகளையும் ஒப்பிட்டு அறிய இயலுகிறது.

(iii) பொருளாதாரப் புள்ளிவிவரங்களைப் பொறுத்த அளவில் பலநாடுகளுடன் ஒப்பீடு செய்வதற்கு, ஐக்கிய நாட்டுப்

புள்ளி ஆய்வுக் கழகத்திற்கு இந்த அமைப்பு மிக்க துணை செய் கிறது. பல வெளியீடுகளுக்கும் உதவி புரிகிறது.

(iv) தேசீய வருமானம் (national income) பற்றிய மதிப்பீடுகளைக் கணக்கிட்டு அறிவதுடன், பொருளாதாரத் திட்டங்கள் அமைத்திட அதிக அளவில் உதவி புரிகிறது.

(v) வெவ்வேறு அமைச்சகங்களுக்கும், அரசுத் துறைகட்கும் தேவைப்படும் அறிவுரைகளைப் புள்ளிவிவர ஆய்வினைப் பொறுத்தவரையில் வழங்கி, பல்வேறு ஆய்வுகளையும் தக்க விதத்தில் தொடர்புடையதாகச் செய்கிறது.

(vi) மக்கட்தொகை, பிறப்பு—இறப்புவீதம் போன்ற பல பிரச்னைகளைப் பற்றிப் பரிசீலித்து ஆய்வுகள் நடத்தி முடிவுகள் எடுக்கப்படுகின்றன.

(vii) பல்வேறு நாடுகளின் வெவ்வேறு கூட்டமைப்புகளுடன் தொடர்ந்து நெருங்கிய தொடர்புகொண்டு, பொருளாதார முன்னேற்றம், ஆய்வு முறைகள், திட்ட அமைப்பு போன்ற பலவற்றில் சிறந்த கவனத்துடன் பணி புரிந்து, நாட்டின் இன்றியமையாத உறுப்பாக விளங்குகிறது.

NSS க்குத் தேவையான பல்வகையான புள்ளி விவரங்களை, CSO வழங்குகிறது. இதன் உதவியின்றி, எந்தவிதமான புள்ளி விவர ஆய்வினையும் முழு அளவில் மேற்கொள்வது கடினம். மேலும், (i) இந்திய நாட்டின் புள்ளிவிவரத் தொகுப்பு (Statistical Abstract of Indian Union) ; (ii) தொழில் துறை பற்றிய வருடாந்திர ஆய்வுகள் (Annual Survey of Industries) ; (iii) புள்ளிவிவர மாதாந்திரத் தொகுப்பு வெளியீடு (Monthly Abstract of Statistics) ; (iv) தேசீய வருமான மதிப்பீடுகள் (National Income Estimates) போன்ற பலவகையான இதன் வெளியீடுகள் பெரும் அளவில் உதவிபுரிகின்றன.

NSS க்குக் கருதப்படும் பொருள் (subject) மற்றும் ஆய்வுவழி முறைகள் (methodology) போன்றவற்றின் திட்ட வரையறையினை CSO அளிக்கிறது. அதன்படி NSS மேற்கொள்ளப்பட்டு, ஆய்வுகள் செம்மையாக நடத்தி முடிக்கப்பட்டு, தேவையான முடிவுகள் பெறப்படுகின்றன.

இனி, கல்கத்தாவில் இயங்கி வரும் ISI என்னும் நிறுவனம் பற்றிய ஒரு குறிப்பினைக் காண்போம்:

4. இந்தியப் புள்ளிவிவர ஆய்வு நிறுவனம் (1932) (Indian Statistical Institute)

சுருக்கமாக 'ISI' என்று குறிப்பிடப்பெறும் இந் நிறுவனமானது அரசினரால் அமைக்கப்படாத, தனியார் துறையினைச் சார்ந்த கூட்டமைப்பாகும். இது தற்போது ஒரு சிறந்த வல்லமை பெற்ற அமைப்பாக விளங்கி வருகிறது. புள்ளியியலின் பல்வேறு பயன்களைப் பற்றிப் பலவகையாக விரிவாய்வுகள் நடத்திச் சிறந்த முடிவுகளை இந் நிறுவனமானது பல துறைகட்குப் பயன்தரக் கூடிய விதத்தில் அளிக்கின்றது. பல பிரச்சினைகளைச் சிறந்த வழிமுறைகள்மூலம் தீர்க்கவும் செய்கிறது. திட்டமிடுதல்பற்றிய ஒரு பிரிவானது இதில் விளங்கி, திட்டக் குழுவிற்கு (மத்திய அரசினைச் சார்ந்தது) மிகவும் உதவி புரிகிறது. அவ்வப்போது இது பலவிதமான ஆராய்ச்சிகளின்மூலம், மொத்த விளைச்சல் மதிப்பீடு, உற்பத்தி அளவு போன்ற பலவகை மதிப்பீடுகளைப் பெறுவது. பல்வேறு நிலைகளில் பெரிதும் உதவிபுரிவதைக் காணலாம்.

NSS ஆய்வுகளைப் பொறுத்தமட்டில், ISI-யின் பணியானது மிகப் பரந்த அளவில் அமைவதொன்றாகும். ஆய்வின் திட்டமிடுதல், மாதிரி வடிவ அமைப்பு, வினாப்பட்டியல்களைத் தயாரிப்பது, தக்க குறிப்புகளை (instructions) இணைப்பது, களப் பணியாளர்கட்குத் தேவையான பயிற்சி அளித்தல், விவரங்களை வழிமுறைப்படுத்தித் தக்க அட்டவணைகளை அமைத்தல், மேலும் இறுதி அறிக்கைகளை விரிவான முறையில் தெளிவாக அமைத்தல், போன்ற விதமாக அமையும் ஆய்வின் பல்வேறு செயல்முறைகள் இந் நிறுவனத்தின் மூலமே நடைபெறுகின்றன. எனவே, NSS ஆய்வுகளில், ISI யின் முக்கியத்துவமானது எளிதில் அறியப்படும். மாதிரி ஆய்வு நடத்தும் கூட்டமைப்பின் (agency) தேவைகட்கு ஏற்றவகையில் தெளிவான திட்டத்தினை உருவாக்கி வழிமுறைகள் வகுக்கப்பட்டு, NSS-ன் மேல்மட்ட ஆய்வுக் குழுவினரால் (Advisory Committee) நன்கு பரிசீலிக்கப்பட்டு, திருத்தங்கள் மேற்கொள்ளப்பட்டு, அதன்பின் களப்பணியானது தக்கபடி துவங்கப்பெறும். எனவே, ISI என்ற கூட்டு அமைப்பானது, NSS ஆய்வுகளில் ஒரு பெரும் இடத்தினைப் பெறுவது என்று கூறுதல் மிகையாகாது.

5. *NSS*-ன் மறு அமைப்பு (*NSSO*) (Reorganisation)

இந்திய அரசின் முடிவுப்படி, 1970-ம் ஆண்டு, தேசீய மாதிரிக் கணிப்பாய்வு அமைப்பு மறு பரிசீலனை செய்யப்பட்டு

புதியதோர் உருவினைப் பெற்று விளங்குகிறது. அதன்படி, 1971-ல் (இந்திய) தேசீய மாதிரிக் கணிப்பாய்வுக் கழக அமைப்பு (National Sample Survey Organisation) உருவாக்கப்பட்டது. *NSS* நிர்வாகக் குழுவானது (Directorate) தொடர்ந்து செயல் பட்டு வருகிறது. அதன் பணிகள் வரையறுக்கப்பட்டு, கள ஆய்வுப் பணிகளுடன்மட்டும் தொடர்புள்ளதாகத் திருத்தி அமைக்கப்பட்டது. களப் பணிப் பிரிவாக (Field Operations Division) இயங்கி வருகிறது. மேலும், *ISI*-உடன் முழு அளவில் தொடர்பு இல்லாத தனித்த கூட்டமைப்பாக *NSSO*-வினை ஆக்குவதற்கு முயற்சிகள் எடுக்கப்பட்டு வருகின்றன அதன்படி, சில செயல்முறைத் திட்டங்கள் (techniques), அதாவது ஆய்வுத் திட்டமிடல், விவரங்களை வழிமுறைப்படுத்தித் தொகுத்தல் (processing), அறிக்கை தயாரித்தல் போன்ற பணிகள் (கடமைகள்) *NSSO*-வின் மூலமாகவே நிறைவேற்றப்பட்டு வருகின்றன என்பது குறிப்பிடத் தக்கது.

எனவே, *NSSO* அமைப்பானது (i) கணிப்பு ஆய்வு, திட்டவரை, மற்றும் விவரங்களின் தீவிர ஆராய்ச்சி; (ii) கள ஆய்வுப் பணிகள்; (iii) விவரங்களை வழிமுறைப் படுத்துதல்; (iv) பொருளாதாரப் பகுப்பாய்வு (Economic analysis) போன்ற முக்கியப் பிரிவுகளைக்கொண்டு விளங்குகிறது. எனவே, நாளடைவில் இந்தக் கூட்டமைப்பு மிகவும் சிறந்த ஒரு இடத்தைப் பெற்று, அதிக முக்கியத்துவம் வாய்ந்த அமைப்பாகத் திகழ்கின்றது.

NSSO-வின் பணிகள், ஒரு மேல்மட்ட ஆளும்-குழுவினரால் (Governing Council) மேற்பார்வை செய்யப்பட்டு, கட்டுப் படுத்தப்படுகின்றன. இத்தகைய குழுவில், அரசினரால் நியமிக்கப்பட்ட பொருளாதார, மற்றும் புள்ளியியல் வல்லுனர் களும், அரசுத்துறை அல்லாத சிலரும் இடம்பெறுகின்றனர். *NSSO*-வின் முக்கிய நிர்வாக அதிகாரி (Chief Executive Officer), இதன் காரியதரிசியாக (செயலாளர்) விளங்குவார்.

ஒவ்வொரு ஆண்டும், *NSSO* ஆனது மத்திய அரசின் நிதி உதவிகளைப் பெறுகிறது. அதற்குத் தக்கபடி, *NSSO*-வின் செயல்முறைத் திட்டமானது (programme) உருப்படுத்தப் படுகிறது. மேலும், விவரங்களைச் சேகரிப்பது, தக்கவழியில் ஆய்வு வழிமுறைகள் வகுப்பது, முடிவுகளை வெளியிடல்போன்ற பல பணிகளை நிறைவேற்றுவதில் *NSSO* அமைப்பானது முழு

அளவில் உரிமை பெற்றுள்ளது. *NSS* சுற்றுககள, மாதிரி உரு அமைப்பு, பட்டியல் அமைத்தல், வழிமுறைப்படுத்தல் போன்ற பல்வேறு செயல்களைப்பற்றி விவாதித்து, தக்க முடிவுகளைப் பெறுவதற்கு ஆளும் குழுவானது (Governing Council) எத்தனை முறை வேண்டுமானாலும் கூடி, விவாதத்தில் ஈடுபடலாம்.

மேலும் *NSSO*-வானது சமூக மற்றும், பொருளாதாரத் தொடர்புடைய பலவகை விவரங்களைத் திரட்டி, பொருளாதாரத் திட்டமிடல் (planning) மற்றும் வேறு சில ஆய்வுகள் ஆகியவற்றுக்கு, மத்திய மற்றும் மாநில மட்டங்களில் அரசினருக்குத் தேவையான அளவில் உதவி புரிகின்றது. பல நோக்கங்கள் கொண்ட, தொடர்ச்சியான, கூட்டுத்தன்மையுடைய ஆய்வுகள் நடத்தப்பட்டு, கருதப்படும் காலத்தில் (current period) முக்கியத்துவம் பெறும் பிரச்சினைகள் தீவிரமாக அலசி ஆராயப்படுகின்றன. 'சுற்றுக்கள்' (rounds) மூலம் ஆய்வு முறைகள் மேற்கொள்ளப்படுகின்றன. கிடைக்கும் தகவல்கட்கு (resources) ஏற்பவும், தேவைப்படும் விவரங்களின் இயல்பிற்கு ஏற்பவும் தக்க பொருத்தமான விதத்தில், நோக்கம், ஆய்வுக்காலம், ஆய்வுச் செயல்முறைத் திட்டம் போன்றவை அமைக்கப்படுகின்றன. 14-ஆவது சுற்றிற்கு (round) முன்பு வரையில், 'சுற்றுக் கால' வரையானது (period of the round) 3 முதல் 8 மாத கால அளவினதாகக் கருதப்பட்டு வந்தது, ஆனால், அதன் பின்னர் (அதாவது, 14-வது சுற்றிலிருந்து) ஆய்வுக் காலம் ஓராண்டாகக் கருதப்பட்டு வருகிறது. இது, விவசாயத் துறையில் உள்ள காலவரையோடு (agricultural year) பொருந்தும் வகையில் அமைகிறது.

6. NSS-ல் மாநிலங்களின் பங்கு

நாட்டின் வெவ்வேறு மாநிலங்களைச் சார்ந்த புள்ளி விவர ஆய்வுக் கழகங்கள் (State Statistical Bureaus) NSS ஆய்வுகளில் முக்கியப் பங்கு வகிக்கின்றன. ஏறக்குறைய, கருதப்பட்ட மாதிரியின் மூன்றில் இரு பகுதியில், கள ஆய்வுப் பணியானது மாநிலங்கள் அளவில் மேற்கொள்ளப்பட்டு, ஆய்வின் இறுதி முடிவுகள் நன்கு சரி பார்க்கப்பட ஏதுவாகிறது. குறிப்பாக, மக்களிடையே சொந்த நிலத்தின் அளவுபற்றிய ஆய்வுகளில், மாதிரி அளவினை அதிகப்படுத்துவதில் அக்கறை எடுக்கப்பட்டது. விவரங்களைத் திரட்டுவதில், மத்திய மற்றும் மாநிலங்களைச் சார்ந்த வெவ்வேறு அமைப்புக்கள் ஒரே வகையான வினாப் பட்டியல்களையும், குறிப்புகளையும் (instructions) பயன்படுத்துவதால்,

விவரங்களைச் சரிபார்க்க இயலுகிறது. தற்சமயங்களில், எல்லா மாநிலங்களும், NSS ஆய்வுப் பணிகளில் தகுந்த பங்கேற்று, இறுதிநிலையான அட்டவணை அமைத்தலிலும் தனிக் கவனம் செலுத்துகின்றன. ஒவ்வொரு மாநிலத்துக்கும் மொத்தமாக அமையும் மாதிரியானது, நான்கு (அல்லது அதன் பெருக்கல் அளவு), 'ஒன்றினை ஒன்று ஊடுருவத்தக்க துணை மாதிரிகள் அமைந்த கூருகத் தெரிவு செய்யப்படுகிறது. இவற்றில் இரண்டு (அல்லது அதிகமான) துணைமாதிரிகள் (sub-samples) மாநில அமைப்புகட்கு (state agencies) ஒதுக்கப்பட்டு, மீதம் உள்ளவை மத்தியக் கழகத்தின்மூலம் கருதப்படுகின்றன. மொத்த ஆய்வானது இவ்வாறு பிரித்துக்கொள்ளப்படுவது, முடிவுகளைச் சிறந்த முறையில், நம்பத்தக்க விதத்தில் அடையும் பிரச்சினை எளிதாக்கப்படுகிறது. மேலும், துணைமாதிரிகள் அனைத்தும் வெவ்வேறு ஆய்வாளர்கள்மூலம் தனித் தனியாக ஆய்வு செய்யப்படுகின்றன. மத்திய மற்றும் மாநில அமைப்புக்களின் இத்தகைய கூட்டு ஆய்வுப் பணி, NSS-ன் குறிப்பிடத்தக்க அம்சமாக அமைந்திருக்கிறது. இறுதி முடிவுகள், மத்திய அரசின் திட்டங்களுக்கு மட்டுமன்றி, மாநில அரசுகட்கும் பெரிதும் உதவிபுரிகின்றன. எனவே, ஒவ்வொரு மாநிலமும் போட்டி மனப்பான்மையுடன் தம் பணியை நிறைவேற்ற முயலுகிறது.

7. *NSS*-ல் கணிப்பு முறைகள் (Methodology)

நன்கு பயிற்சிபெற்ற, முழு நேர ஆய்வாளர்களைப் பயன் படுத்தித் தேவையான விவரங்கள் அனைத்தும் சேகரிக்கப்படுகின்றன. சமூகம் மற்றும் பொருளாதார (socio-economic) இயல்புகளைப் பற்றிய விவரங்கள் அனைத்தும், மாதிரியில் இடம் பெறும் நபர்களை அல்லது குடும்பங்களை நேரடியாகத் தொடர்பு கொண்டு பெறப்படுகின்றன. பயிர் விளைவு ஆய்வுகளில் (crop surveys) தெரிந்தெடுக்கப்படும் நிலக் கூறுகளில் நேரடிக் கணிப்பு செய்யப்பட்டு, விளைவு வீதமானது கண்டறியப்படுகிறது. பல்வேறு கணிப்புகளில், நிலைமைக்கு ஏற்றவாறு ஒருநாள், ஒரு மாதம், ஓர் ஆண்டு என்றபடி, குறிப்பீட்டுக் காலங்கள் (reference periods) கருதப்படுகின்றன.

பலபொருள் ஆய்வு (multi-subject survey) நடத்துவதன் மூலம், கணிப்பினில் கருதப்படும் பல்வேறு பொருட்களை அல்லது இயல்புகளை ஒரே மாதிரியின் முதல்நிலைக் கூறுகளில் (first stage units) இணைத்துச் செயல்படுவதன்மூலமாகக் கிடைக்கும் மூல

விவரக் கூறுகளை திறமையுடன் பயன்படுத்தவும், ஆய்வு நேரம் மற்றும் ஆகும் செலவு இவற்றைக் குறைத்துப் பயனடையவும் இயலுகிறது. அதாவது, தெரிந்தெடுக்கப்படும் கிராம அல்லது நகரத் தொகுதிகளிலேயே பல்வேறு பொருள்கள் பற்றிய விரிவான கணிப்பாய்வும் அளவெடுத்தலும் நடத்தப்படுகின்றன. இதற்காக, மாதிரி ஆய்வின் முதற் கட்டத்திலேயே (first stage), மாதிரி ஆளவானது (sample size) கணிசமாக உயர்த்தப்படுகிறது. இதன்மூலம் பெறப்படும் மதிப்பீடுகளும் அதிக திட்பம் வாய்ந்தனவாக விளங்குகின்றன. இத்தகைய பல்பொருள் ஆய்வுகள் சிக்கனம் வாய்ந்தன ஆதலால், இம்முறையிலேயே வழிமுறைகள் வகுக்கப்படுகின்றன. திறமையான முடிவுகளையும் பெற முடிகிறது. மேலும், முதல்நிலைக் கூறுகளின் இடைப்பட்ட மாறுபாடானது பல்வேறு இயல்புகட்கும் மிக்க கவனத்துடன் கருதப்படவேண்டியதாக அமைகிறது. பன்முகத் தேர்வு (multiphase) முறைகள் மூலமாக கணிப்பாய்வுகள் நடைபெறுகின்றன எனலாம்.

8. NSS-ஆய்வின் செயல்முறை (Techniques in Surveying) NSSO-வானது செயல்படும் வழிமுறையானது (methodology)

(i) ஆய்வுக் கூறினைத் தேர்ந்தெடுத்தல் (Selection of unit)

(ii) ஆய்வின் திட்ட வரையரை (design)

(iii) களப்பணியினைத் திட்டமிடல் (Field-work)

என மூவகையாகக் கருதப்படலாம்.

(i) முதலில், ஆய்வு நடத்தப்படும் பரப்பில் கூறுகளை அல்லது அலகுகளைத் தக்கவிதத்தில் தேர்ந்தெடுக்கவேண்டும். ஆய்வின் அடிப்படை நோக்கமானது, மக்களின் பொருளாதார மற்றும் சமூக வாழ்வின் பல்வேறு தன்மைகளை விளக்கத்தக்க விவரங்களைத் திரட்டுவதே ஆகும். நம் நாட்டினைப் போன்று மிகவும் பரந்த எல்லைப் பரப்பில் அமையும் நாடுகளில், புள்ளி விவரங்களை ஒன்றுதிரட்டித் தொகுப்பது அவ்வளவு எளிதான செயல் அல்ல. எனவே, வாய்ப்பு அடிப்படையில் மாதிரிகளைத் தெரிவு செய்யும் வழிமுறை கையாளப்பட்டு வருகிறது.

முதன்முதலில், சமவாய்ப்பு மாதிரித் தேர்வு முறையில் (random sampling method), கிராமப்புறங்கள் (villages) தேர்ந்தெடுக்கப்பட்டன. ஆனால், ஆரம்ப கட்டத்தில், கிராமப்புறங்களைப் பற்றித் தெளிவாக விளக்கத்தக்க, கிராமம் மற்றும்

தாசில் (தாலுக்காக்கள்) (tehsil) பற்றிய விளக்கப் படங்கள் (maps) பல இடங்களில் அமைக்கப்படவில்லை. எனவே, நிலப் பரப்பு (area), மற்றும் மக்கட்தொகை (population) ஆகியவை பற்றிய விவரங்கள் (செய்திகள்) கொண்ட விரிவான பட்டியலைக் (list) கருதவேண்டிய நிலை ஏற்பட்டது. இதிலும் சில சிக்கல்கள் ஏற்பட்டன. ஏனெனில், சில பரப்புகளில், மக்கட்தொகை அளவு விவரம் மற்றும் கிராமப்புறங்களின் பெயர்கள் விவரம் ஆகியவை கிடைக்கப்பெருததால். அவற்றினை, களப் பணியாளர்களே (field staff) தக்க ஆய்வுகள்மூலம் கண்டறிய வேண்டியதாயிற்று.

(ii) NSS—திட்ட அமைப்பு (Design)

NSS செயல் முறையின் விரிவான திட்டமானது (scheme), பி. ஸி. மஹலநோபிஸ் (P.C Mahalanobis) அவர்களால் வகுக்கப்பட்டு உருவாக்கப்பட்டதாகும். ஆய்வின் வடிவ அமைப்பினை, ISI எனப்படும் இந்தியப் புள்ளியியல் கழக மானது உருவாக்கித் தந்தது. இதன்மூலம், ஆய்வு நடத்தப் பெறும்விதம், செயல்முறைத் திட்டம், வினுப் பட்டியல்களையும், வினுத் தொகுப்புக்களையும் தயாரித்தல், ஆய்வுச்செலவு மதிப்பீடு, இறுதி வடிவம் பெறல் ஆகிய பல்வேறு பணிகளையும் ISI மேற்கொண்டு, NSS-ன் செயல்முறைத் திட்டத்தினைத் தெளிவாக உருவாக்கிற்று. இந்த வழியிலேயே ஆய்வுகள் அனைத்தும் (NSS) நடத்தப்பட்டு, தேவைப்படும் முடிவுகள் எந்தவிதத் தடங்கலும் இன்றிப் பெறப்படுகின்றன. ஆய்வு நோக்கத்தின் மிக இன்றியமையாததாகும் இது. சிக்கலற்ற செயல்முறையை வகுப்பது களப்பணியை (field work) எளிதாக்கி விரைவில் முடிவுகளைப் பெறுவது சாத்தியமாகிறது.

(iii) களப்பணியின் உருவ அமைப்பு (Field-work)

இனி கள ஆய்வுப் பணியின் திட்டமிடலைக் கருதுவோம். அலுவலகப் பணிகள் அனைத்தும் (office work), ISI மூலமாக நடைபெறுகின்றன. கள ஆய்வுப் பணியானது, மத்திய நிதி அமைச்சகத்தின் பொருளாதார விவகாரங்கள்பற்றிய இலாகா வினைச் சார்ந்ததாகக் கருதப்படுகிறது. வெவ்வேறு மாநில அரசு களும், கள ஆய்வுப் பணிக்குத் தேவையான அலுவலர்களைத் (officers) தருமாறு கேட்டுக்கொள்ளப்படுகின்றன. துணை இயக்குநர்கள் (Asst. Directors), மற்றும் கண்காணிப்பாளர் (அல்லது மேற்பார்வையாளர்) (Superintendents) ஆகியோர் மேல் மட்டத்தில் மாநில அரசுகளால் நியமிக்கப்படுவர். சிலத்

நேரடித் தேர்வு பெறுவதும் உண்டு. கள ஆய்வாளர்கள் மற்றும் கணக்கெடுப்போர் (Enumerators) ஆகியோரும் நியமிக்கப்பட ஆய்வுப் பணிக்கு தேவையான அலுவல் குழுவானது முழு அமைப்பாக உருவாகிறது. இது களச் செயல் முறையின் முதற் கட்டமாகும்.

அடுத்த கட்டமான பயிற்சி அளிக்கும் நிலையில் (training), நியமிக்கப்பட்ட வெவ்வேறு வகையான அலுவலர்கட்கும் மூன்று வார கால அளவினதான பயிற்சி அளிக்கப்பட்டது. மேல்மட்ட அதிகாரிகள்மூலம், கீழ்நிலை ஆய்வாளர்கட்கு (investigators) விவரங்களைத் திரட்டும் முறைகள்பற்றிய விரிவான மற்றும் விளக்கமான பயிற்சி அளிக்கப்படுகிறது. நன்கு பயிற்சிபெற்ற மேலதிகாரிகள், இந்தியாவில் ஆறு தெரிந்தெடுகப்பட்ட மையங் களில் (centres), கள ஆய்வுப் பணியாளர்கட்குத் தேவையான பயிற்சியினை அளிக்கின்றனர்; இப்பயிற்சியோடுகூட, விரிவான குறிப்புகள்கொண்ட கை ஏடு (hand-book) ஒன்றும் ஒவ்வொரு ஆய்வாளருக்கும் அளிக்கப்படுகிறது. இதன்மூலம் வெவ்வேறு மட்டங்களில் கள ஆய்வில் ஈடுபடும் அலுவலர்கள் அனைவரும் நேரடிப் பணித் துவக்கத்திற்கு ஆயத்தப்படுத்தப்படுகின்றனர்.

முதல் நிலை முறையான (primary method), நேரடிக் கணிப்பு முறையின் (personal investigation) மூலமாக, கணக்கெடுப்போர் தேவையான விவரங்களைத் திரட்டுகின்றனர். தெரிவு செய்யப் படும் நபர்களை நேரடியாக அணுகி, தேவைப்படும் விவரங்கள் அனைத்தையும் பெறுவது அவர்தம் கடமையாக வகுக்கப்படுகிறது. கண்டறியப்பட்ட விவரங்களின் உறுதிப்பாட்டினை, மீண்டும் மீண்டும் ஆராய்ந்து உறுதிப்படுத்திக்கொள்ளுதல் இன்றியமை யாததாகும். விவரங்கள் பொருத்தமாகவும், இயன்றவரையில் துல்லியமான (accurate) அளவுடையனவாகவும் இருக்க வேண்டும். அவர்தம் தீவிரமான பயிற்சியின்மூலம், பெறப்படும் மதிப்பீடுகள் அனைத்துமே நம்பகத்தன்மை உடையனவாகக் கருதப்படும். குறிப்பாக, நிலங்களின் பயன்பற்றிய ஆய்வு களிலும், பயிர் சாகுபடிபற்றிய மதிப்பீடுகளிலும், ஆய்வாளரது முடிவே ஏற்றுக்கொள்ளப்படும். பயிர்களின் சூழ்நிலைபற்றிய விவரங்களையும், பயிர் செய்யப்படும் மொத்தப் பரப்பின் தன்மை களையும்பற்றித் தெளிவாக அறிந்தவராக ஆய்வாளர் விளங்குவது ஏற்கத்தக்க முடிவுகளைப் பெற உதவுகிறது. தேர்ந்தெடுக்கப் படும் ஆய்வாளர்களில் பெரும்பாலோர் கிராமப்புறங்களினின்று வருவதாக அமைவதால், ஆய்வில் பெறப்படும் மதிப்பீடுகள் சிறந்து விளங்க ஏதுவாகிறது. மேலும், பெறப்பட்ட விவரங்

களும், மதிப்பீடுகளும் மேல்நிலைக் கண்காணிப்பாளர்களால் தக்கபடி சரிபார்க்கப்படுகின்றன.

9 NSS-ஆய்வின் பொதுவான செயல் திட்டம் (General Plan)

நாட்டின் மொத்த நிலப் பரப்பினையும் கருதி, கள ஆய்வின் மூலமாக, விவரங்களைச் சேகரிப்பதற்கு உதவும் பொருட்டு, மொத்தப் பரப்பும் 16 (தொகுப்பு) பிரிவுகளாக (blocks) பகுக்கப் படுகிறது. ஒவ்வொரு பிரிவும் அல்லது தொகுப்பும், ஏறக்குறைய ஒரு மாநிலத்தின் அளவு பரப்புடையதாகக் கொள்ளப்படும். சில நிலைகளில், அளவின் தன்மை கருதி, இரண்டு அல்லது மூன்று மாநிலங்களை இணைத்து ஒரு தொகுப்பானது கருதப்படுகிறது. உதாரணமாக பரப்பளவில் பெரிதாக அமையும் உத்தரப் பிரதேசம் போன்ற மாநிலங்கள் தனியான தொகுப்பாகக் கருதப் படுகின்றன. இதற்கு மாறாக, அசாம், திரிபுரா, மணிப்பூர் போன்ற சிறு சிறு அளவினதான மாநிலங்களை இணைத்து தொகுப்பானது உருவாக்கப்படுகிறது. ஒவ்வொரு தொகுப்பி லிருந்தும், ஆய்வுக்காகச் சில கிராமங்கள் தேர்ந்தெடுக்கப் படுகின்றன. நன்கு பயிற்சிபெற்ற முழு நேர ஆய்வாளர்கள் ஒவ்வொரு கிராமத்திற்கும் நேரடியாகச் சென்று, மக்களிடம் விவரங்களைப் பெற்று அறிந்து, பட்டியலில் குறித்துக்கொள் கின்றனர். விவசாயத்துறை பற்றிய விவரங்களை, ஆய்வாளர்கள் தாமே நிலங்களைப் பார்வையிடுதல்மூலமாக மதிப்பிட்டு அறிகின்றனர்.

இந்திய நாட்டின் பரப்பினை முழுதும் காட்டும் வரைபடத்தின் (map) அடிப்படையில், ஒவ்வொரு மாநிலத்திலும், மாநில, மாவட்ட மற்றும் வட்ட அளவுகளில் எல்லைகள் (boundaries) சரிவர நிர்ணயிக்கப்படுகின்றன. கடல் மட்டத்துக்குமேல் உள்ள உயரத்தினைக் கருதி அதன் அடிப்படையில் ஒவ்வொரு மாநிலத்தின் வட்டப் பகுதிகளும் மூன்று படுகைகளாகப் பிரிக்கப் படுகின்றன. அதன்பின்னர் மக்கள் நெருக்க (Population density) அடிப்படையில், மீண்டும் படுகைகள் பிரிக்கப்படும். பெரிய மாநிலங்களை (உத்தரப் பிரதேசம் போன்றவை) பொறுத்த மட்டில், முக்கிய விளைபொருள்களைக் கருதி, அதற்கேற்ப இரண்டு அல்லது மூன்று பிரிவுகளாகக் கொடுக்கப்பட்டு, உணவுப் பயிர் வகைகள்பற்றிய விவரங்கள் பெறப்படும். ஆய்வின் எளிமை கருதி இவ்வாறு மேற்கொள்ளப்படுகிறது. தாசில்களை அல்லது வட்டங்களைத் தக்கபடி அமைத்துக்கொண்ட பிறகு தொடர்ச்சி யாக உள்ளவற்றை இணைத்து, தேவையான படுகைகள்

அமைக்கப்படுகின்றன. படுகைத் தொகுதிகள் சமமாக அமையுமாறு கவனித்துக்கொள்ளப்படுகிறது.

வெவ்வேறு இடப் பரப்புகளின் தன்மைகளையும், காணப்படும் தரங்களையும் (standards) பிரதிபலித்துக் காட்டத்தக்க விதத்தில், மாதிரி கிராமங்களைச் சமவாய்ப்பு நியதி அடிப்படையில் தெரிவு செய்து விரிவான ஆய்வு நடத்தப்படுவதே பொதுவான செயல்முறையாகும். ஒவ்வொரு சுற்றிலும் (round) பெறும் அனுபவத்துக்கும், விளையும் இடர்ப்பாடுகட்கும், கண்டறியப்படும் முடிவுகட்கும் ஏற்றபடியாக, கிராமங்கள் தெரிந்தெடுக்கப்படும் முறையில் மாற்றங்கள் செய்யப்படுகின்றன. அடுத்த பகுதியில், 'மாதிரித் தேர்வு' பற்றிய கருத்துக்களில் சமீபகால சுற்றுகளில் NSS ஆய்வுகளில் மேற்கொள்ளப்படும் வழிமுறைகள் விளக்கப்படுகின்றன. கால வளர்ச்சிக்கு ஏற்ப, கிராமங்களின் எண்ணிக்கை அதிகரிப்பதால், இத்தகைய மாற்றங்கள் அவசியமாகின்றன.

பலகட்ட மாதிரி முறை (multi-stage sampling) மற்றும் பல்தோற்ற மாதிரி முறை (multi-phase) ஆகிய இருமுறைகளும், பெரும்பாலும் NSS ஆய்வுகளில் பயன்படுத்தப்படுகின்றன. முதல் கட்டத்தில், தேவைப்படும் அளவில் கிராமங்கள் தெரிவு செய்யப்படும். அடுத்து விரிவான ஆய்வு நடத்துதற்பொருட்டு, சமவாய்ப்பு அடிப்படையில், குடும்பங்கள் தெரிவு செய்யப்பட்டு, பல நோக்கங்கள் கொண்ட ஆய்வுகள் நடத்தப்படும். வெவ்வேறு இயல்புகளைப் பற்றி ஒரே தன்மையான துணைக்கூறுகளைக் (அதாவது குடும்பங்களை) கருதி, விரிவான ஆய்வு மேற்கொள்ளப்படும்.

கிராமப்புற நிலங்களைப் பற்றிய விவரங்கள் சேகரிப்பது, 'நிலத்தைப் பயன்படுத்தும் முறை' (land utilization) பற்றிய ஆய்வுக்குப் பெரிதும் துணைபுரிகிறது. துவக்க நிலைகளில் ஆய்வுகள் பெரும்பாலும் கிராமங்களைக் குறித்தே அமைந்தன. மூன்றுவது சுற்றிலிருந்து, ஆய்வு நோக்கமானது நகர்ப்புறங்கட்கும் விரிவாக ஆக்கப்பட்டு, ஆய்வுத் தன்மையும் பரந்த அளவினதாக அமைந்தது. பின்வரும் சில கருத்துக்கள் மிகச் சமீப கால ஆய்வுகளை (1971 மக்கட்தொகைக் கணிப்பினைச் சார்ந்து) அடிப்படையாகக் கொண்டவையாகும்

10. அறிவிப்புக் காலம் (Reporting Period)

இந்திய நாட்டின் பொருளாதார அமைப்பானது பெரும்பாலும் விவசாயத் துறையினையே பற்றி அமைந்

திருப்பதால், கிராமப்புறங்களில், பருவகால மாறுபாடுகளால் (seasonal variation), ஆய்வு நடத்துவதில் சில சிரமங்கள் ஏற்படுகின்றன. பொதுவாக, ஆய்வுக்காலவரையானது (survey period), ஓராண்டாக நிர்ணயிக்கப்படுகிறது. இத்தகைய முடிவானது, *NSS*-ன் 14-வது சுற்றில் எடுக்கப்பட்டு. இன்றுவரையில் கடைபிடிக்கப்பட்டு வருகிறது. மேலும், ஒரு சுற்றுக்கும் அடுத்த சுற்றுக்கும் இடையிலுள்ள கால அளவினையும் மிகவும் குறைப்பதற்கு வகை செய்யப்பட்டுள்ளது தெரிந்தெடுக்கபட்ட குடும்பங்களிலிருந்து விவரங்களைச் சேகரிப்பதில், ஆய்வு நடத்தப்பெறும் காலத்திற்கு ஒருநாள் அல்லது ஒருவாரம், அல்லது ஒருமாதம் முன்பாகவே 'அறிவிக்கும் காலம்' நகரும் தன்மையதாகக் கருதப்பட்டு, தேவையான விவரங்கள் பெறப்படுகின்றன. விவரங்கள் அறிவிக்கப்படுவதற்கேற்ப அல்லது தெரிவிக்கப்படுவதற்கேற்ப, ஆய்வு நடைபெறும் குறித்த காலத்தினில் வெவ்வேறு இயல்புகளின் சராசரிகள்பற்றிய மதிப்பீடுகளைப் பெறமுடிகிறது. எனவே ஒரு குறிப்பிட்ட நேரத்தைக் கருதி ஆய்வு நடத்தப்படுவதுதிலிருந்து இம்முறை வேறுபடுகிறது. பயிர்விளைவுபற்றிய ஆய்வுகளில் (crop surveys) வெவ்வேறு பயிர்கட்கு வெவ்வேறு பருவகாலம் அமையப் பெறுவதால், எல்லாவற்றுக்கும் பொதுவான ஒரு குறிப்பீட்டுக் காலத்தினை (point of time) எடுத்துக்கொள்வது இயலாத காரியம். எனவே அறிவிப்புக் காலமானது, குறிப்பாக விவசாய நிலைபற்றிய ஆய்வுகளில், நகரும் தன்மையதாக (moving) கருதப்படுவது நன்று.

எப்போதும் ஓரிடத்தில் நிலையாக நில்லாது, இடம் விட்டு இடம் செல்லும் கள ஆய்வாளர்கள் *NSS*-ல் பயன்படுத்தப்படுவதால்,ஒரு குறிப்பிட்ட அறிவிப்புக்கால அடிப்படையில் செயல்பட்டு விவரங்களைச் சேகரித்தல் நடைமுறைக்கு ஒவ்வாததாகும். எனவே, பொதுவாக *NSS* ஆய்வுகளில், ஆய்வுக்காலத்தினை மிகவும் பரந்த அளவில் (wide spread) கருதுகின்றனர். எனவே அறிவிப்புக் காலம் கடந்து, சிறிது காலத்திற்கு பின் ஆய்வானது நடத்தப்படுகையில், அத்தகைய கூறுகளில், சில விவரங்களை மீண்டும் நினைவுகூர்வதில் ஏற்படும் பிறழ்ச்சி அதிக அளவில் காணப்படும்.

11. நடைமுறையில் மாதிரி

ஆய்வுகள் அனைத்தும் ஒன்றன்பின் ஒன்று அமையும் வண்ணமாகத் தொடர்ந்து நடத்தப்படுகின்றன. நாட்டில் உள்ள மொத்தம் 560000 (ஏறக்குறைய) கிராமங்களிலிருந்து.

1833 கிராமங்கள் மாதிரியாகத் தேர்ந்து எடுக்கப்படுகின்றன இவை அனைத்தும், நாடு முழுதும் பரந்த அளவில் அமையுமாறு எடுக்கப்படுகின்றன. இதற்குச் சமவாய்ப்பு முறை (random sampling) பயன்படுகிறது. அதாவது, தெரிவுசெய்யப்பட்ட கிராமப்புரங்கள் அனைத்தும் நாட்டின் மொத்தப் பரப்பிலும் பரவலாக (scattered) அமையக் காணலாம். இவ்வாறு தேர்ந் தெடுக்கப்படும் சில பகுதிகள், நெருங்கமுடியாதவாறும் அமைய வாய்ப்பு உண்டு. அதாவது, கொடியவிலங்குகளால் அபாயம் விளையதக்கதாகவும், சூழ்நிலைகள் ஏற்றனவாக இல்லாமலும் சில பகுதிகள் அமைய நேரிடும்.

இவ்வாறு தெரிவுசெய்யப்பட்ட 1833 கிராமங்களும் மேலும் இரு பிரிவுகளாகப் பகுக்கப்படும். அவையும் பரவலாக அமைந்திருக்கும். இப்பிரிவுகள் இரண்டிலும் வெவ்வேறு வகையான வினாப்பட்டியல்கள் (schedules) பயன்படுத்தப் படுகின்றன. முதல் பிரிவில் 1189 கிராமங்கள் கருதப்பட்டு, அவற்றுக்கேற்ற பட்டியல்கள் *ISI* மூலமாகவும், இரண்டாவது பிரிவில் 644 கிராமங்கள் கருதப்பட்டு, அவற்றுற்கேற்ற பட்டியல் கோகலே நிறுவனம் மூலமும் (Gokhale Institute, Poona) தயாரிக்கப்பட்டுப் பெறப்பட்டன. இதில், கிராமங்கள் தெரிவு செய்யப்படும் முறையானது சுவைமிக்கதாகவும் (interesting) அதே சமயம் பின்பற்றத்தக்கதாகவும் (instructive) அமைந் திருக்கக் காணலாம். முன்புபோல், நாட்டின் மொத்த (பூகோள) பரப்பும் 250 படுகைகளாக (பூகோள அடிப்படையில்), வெவ்வேறு பொருளாதார, சமூக வட்டாரச் சூழ்நிலை போன்ற பல இயல்புகளைப் பிரதிபலிக்கத்தக்க விதத்தில் பிரிக்கப்பட்டன. ஒவ்வொரு படுகையிலும், கருதப்படும் கிராமங்களின் எண்ணிக்கையானது மூன்றால் (3) வகுபடும் தன்மையதாகக் கருதி நிர்ணயிக்கப்பட்டது. இதன்மூலமாக, அவற்றை 2 : 1 என்ற விகிதத்தில் பிரிக்க ஏதுவாகிறது.

முதல் பிரிவில் அடங்கிய 1189 மாதிரிக் கிராமங்களில் பின்வரும் விவரப் பட்டியல்கள் (schedules) பயன்படுத்தப் பட்டன :

1. **கிராமப் பட்டியல் (Village schedule) :** இப்பட்டிய லானது, ஏதேனும் ஒரு மாதிரி கிராமத்தில் இடம்பெறும் குடும்பங்கள் அல்லது நபர்களை வரிசைப்படுத்தி அறிதற் பொருட்டு (listing) தேவைப்படுவதாகும். இதன்மூலமாக (i) நிலம் வைத்திருப்போர் விவரம், (ii) தெரிவு செய்யப்பட்ட

சில பொருட்களின் விலைகள்பற்றிய விவரம், (iii) வெவ்வேறு வகையான திறமைபெற்ற (skilled) மற்றும் திறமைபெருத (unskilled) பணியாளர்கள் பெறும் தினசரிக் கூலிவீதம் பற்றிய விவரம் போன்ற சில விவரங்களை அறியவேண்டியிருக்கும்.

2. குடும்பப் பட்டியல் (முதல்வகை) (Household schedule) இத்தகைய பட்டியலானது, பின்வரும் விவரங்களைப் பெறுவதற்காகத் தேவைபடுகிறது :

(i) குடும்ப நபர்களைப்பற்றிய பொது விவரங்களையும் (demographic), பொருளாதார விவரங்களையும் பொறுத்து அமையும் நிலைகளை அறிதல். (உதாரணமாக, வயது, இனம், திருமணம் பற்றியவிவரம் பொருளாதார மற்றும் பணிசெய்யும் நிலை போன்ற சில விவரங்களைக் கருதலாம்)

(ii) நில உடைமையும் (holding of land), நிலத்தை பயன்படுத்தும் வெவ்வேறு முறைகளும் பற்றிய விவரங்கள்

3. குடும்பப்பட்டியல் (இரண்டாவது வகை) : சிறு அளவில் கருதப்படும் குடும்பங்களிடமிருந்து. வீடுகளில் மேற்கொள்ளப் படும் சிறு தொழில்கள்பற்றிய விவரங்களையும், (household enterprises) மேலும்

(i) விவசாயம், கால்நடை வளர்ச்சி; (ii) பணிகள், வணிகம், தொழில், சிறுவினைத் தொழில்கள் (industry crafts) போன்ற பல விவரங்களையும் அறிதற்பொருட்டு இத்தகைய பட்டியல்கள் தேவைப்படுகின்றன. கைத்தொழில்கள், சிறிய அளவில் அமையும் குடிசைத்தொழில்கள், மற்றும் பயிர்ச்சாகுபடி போன்றவை கிராமப் பொருளாதார நிலையின் முக்கியக் கூறு களாக விளங்குகின்றன.

4. குடும்பப் பட்டியல் (நாலாம் வகை) : இத்தகைய பட்டியலும், சிறு அளவில் கருதப்பட்டு அமையும் குடும்பங்களி லிருந்து விவரங்களைப் பெறுவதற்குத் தேவைப்படுவதாகும். இதன் மூலம் (i) உணவு வகைகள் (ii) எரிபொருள், ஒளிவசதி; (iii) வீட்டுவாடகை; (iv) துணிவகை (v) மற்ற தேவைகள் ஆகிய பலவற்றின் தேவைப்படும் அளவு பற்றியும், மதிப்புப் பற்றியும் விவரங்கள் பெறப்படுகின்றன. அதாவது, சாதாரண மக்களின் அன்றாட வாழ்வுக்குத் தேவையான பல்வேறு அம்சங்களையும், அவற்றின் பூர்த்தி செய்யும் தன்மையையும் அறிய இப்பட்டியல் மிக உதவுகிறது.

12. **விவரங்களைப் பெறுவதற்குக் கையாளப்படும் வழிமுறை :**

நாட்டின் வெவ்வேறு பகுதிகளில், வெவ்வேறு வகையான விவரங்கள் கிடைப்பதாக இருக்கும் சூழ்நிலையில், மாதிரிக் கூறுகளைத் தெரிவுசெய்வதற்கு வெவ்வேறு வழிமுறைகள் பயன்படுத்தப்படுகின்றன. அதாவது, எல்லாப் பகுதிகட்கும் பொருந்தும் வகையில் ஒரு பொதுவான வழிமுறையினை வகுப்பது எளிதான காரியமல்ல. பொதுவாக, மாதிரிக் கூறுகளைப் (sampling units) பின்வரும் மூன்று கட்டங்களில் தேர்ந்தெடுக்கும் முறையானது பின்பற்றப்படுகிறது:

(i) முதற் கட்டத்தில், படுகைகள் (strata) பிரிக்கப்படுகின்றன. பூகோள அமைப்பிற்கும் (geographical) இயற்கை அமைப்பிற்கும் ஏற்ற வகையில் எல்லா மாநிலங்களும் மொத்தம் 160 படுகைகளாகப் பிரிக்கப்படுகின்றன. அதாவது படுகைகள் அனைத்தும் ஓரினத் தன்மை (homogeneity) உடையனவாக அமையத் தக்கவிதத்தில் உருவாக்கப்படுகின்றன. அதன்பின், ஒவ்வொரு படுகைப்பரப்பும் நான்கு துணைப் படுகைகளாகப் (substrata) பகுக்கப்பட்டு, பின்வருமாறு மக்கட் தொகை அடிப்படையில் கொடுக்கப்படும். இத்தகைய படுகைகளில் இடம்பெறும் கிராமங்களின் மக்கட் தொகை பற்றிய விவரங்கள் அறியப்பட்டிருத்தல் அவசியமானது.

துணைப் படுகை		மக்கட் தொகை
முதலாவது	...	1 முதல் 499 வரை
இரண்டாவது	...	500 முதல் 99 வரை
மூன்றாவது	...	1000 முதல் 1999 வரை
நான்காவது	...	2000 - க்கு மேல்

மக்கட் தொகை விவரங்கள் தெரிவதற்கு ஏற்ப, 32 படுகைகளில் மட்டுமே துணைப் படுகைகளை உருவாக்க முடிந்தது என்பது குறிப்பிடத்தக்கதாகும். இவை அனைத்தும் இறுதியில் 128 இறுதிப் படுகைளாகப் பிரிக்கப்பட்டன. எனவே மொத்தமாக 256 படுகைகள் அமையப்பெற்றன. இவற்றிலிருந்து மாதிரி கிராமங்கள் தெரிவுசெய்யப்பட்டன. அது பின்வரும் அடுத்த கட்டத்தில் (stage) கருதப்படுகின்றது.

ஒவ்வொரு படுகையினின்றும் கிடைக்கும் விவரங்களின், தன்மையைப் பொறுத்து, கிராமங்கள் தேர்ந்து எடுக்கப்பட்டு

ஆய்வின் அடிப்படையானது அமைக்கப்பட்டது. இவ்வாறு தெரிவு செய்தற்கு, குறிப்பாக இக் கட்டத்தில் ஐந்து வழிமுறைகள் பயன்படுத்தப்படுகின்றன :

(i) ஒவ்வொரு கிராமத்திற்கும் தெரிவு செய்யப்படுதற்கு சம வாய்ப்பு அளவுடன் (equal probability) அமையும் முறை.

(ii) கிராமத்தின் பரப்பளவு (area) மதிப்புடன் வீதத் தொடர்புடைய நிகழ்தகவு பயன்படுத்தப்படும் முறை (Proabality Proportional to area).

(iii) கிராமத்தின் மொத்த மக்கட் தொகை அளவுடன் வீதத் தொடர்பு (proportional) உடைய வாய்ப்புடன் கூடிய வழிமுறை (probability proportional to population).

(iv) கிராமப் பரப்பளவுடன். தனித்தனியாக ஒவ்வொரு துணைப் படுகையிலும் கருதப்பட்டு, வீதத் தொடர்புடன் வாய்ப்பு நிலை அமைந்திருக்கும் முறை.

(v) அதேபோல, கிராமத்தின் மக்கட் தொகை அளவோடு ஒவ்வொரு துணைப்படுகையிலும் தனித்தனியாக, வீதத் தொடர்புடைய வாய்ப்புநிலை பயன்படுத்தப்படும் முறை.

இருக்கும் சூழ்நிலைக்குத் தக்கவிதத்தில், மாதிரி முறையானது தேர்ந்தெடுக்கப்பட்டு கையாளப்படுகிறது. பொதுவாக, எளிய சமவாய்ப்புமுறை, மற்றும் PPS முறைத் தேர்வு ஆகிய இருவழி முறைகளும், இக் கட்டத்தினைப் பொறுத்த அளவில், பெரும் அளவினில் பயன்படுத்தப்படுவதைக் காணலாம். கிராமங்களை இரண்டாவது கட்டக் கூறுகளாகத் தெரிவுசெய்து முடித்தபின் அடுத்த நிலை வருமாறு:.

மூன்றுவது கட்டத் தேர்வில், ஆய்வு நடத்தப்படும் கூறுகளான குடும்பங்கள் (households) தெரிவு செய்யப்படுகின்றன. இது ஒரு பன்முகத் தேர்வுஆய்வாக (multi-phase) அமைவது குறிப்பிடத்தக்கது. அதாவது வெவ்வேறு நிலைகளில் குடும்பங்கள் தெரிவு செய்யப்பட்டு, ஆய்வுக்குக் கருதப்பட்டன. பின்வரும் முகங்கள் (phases) குறிப்பிடத்தக்கனவாக அமைகின்றன :

முதலில், தேர்வு பெற்ற ஒவ்வொரு கிராமத்திலும், எல்லாக் குடும்பங்களையும் எடுத்துக் கொண்டு கணக்கெடுப்பு (enumeration) நடத்தப்பட்டது. இது பொதுவான விவரங்களைப் பெற வழி செய்கிறது.

இதன்பின்னர், நபர்களின் பணி அல்லது அலுவல் (occupation) பற்றிய விவரங்கள் அறியப்படுதற்கு, 80 குடும்பங்கள் சமவாய்ப்பு அடிப்படையில் மாதிரியாகத் தெரிந்தெடுக்கப்பட்டன. இதன்மூலம் மக்கள் ஈடுபட்டிருக்கும் பல்வேறு பணிகள்பற்றிய மதிப்பீடு பெறப்பட்டது. இந்த 80 குடும்பங்களுமே, இருவகைகளாகப் பிரிக்கப்பட்டன. அதாவது (i) விவசாயத்தில் ஈடுபட்டிருப்போர், (ii) விவசாயத் தொடர்பு அல்லாத பிற பணிகளில் ஈடுபட்டோர். இது கிராமப்புறச் சூழ்நிலைக்கு ஏற்ற பாகுபாடாகும். இவ்விரண்டு பிரிவுகளினின்றும், ஒவ்வொன்றிலிருந்து 8 குடும்பங்களாக மொத்தம் 16 குடும்பங்கள் (அல்லது இல்லங்கள்) வாய்ப்பு நியதியில் தேர்ந்து எடுக்கப்பட்டன. இந்தப் பதினாறு குடும்பங்களிலிருந்தும், பொதுவான பல விவரங்கள் (general particulars) திரட்டப்பட்டன.

அடுத்து, தெரிவுசெய்யப்பட்ட 16 குடும்பங்களினின்றும் மொத்தமாக 2 விவசாயத் தொடர்புடைய குடும்பங்களையும், 3 விவசாயம் அல்லாத துறைகளில் பணிபுரியும் குடும்பங்களையும் சமவாய்ப்பின் அடிப்படையில் தெரிவுசெய்து, வீடுகளில் நடைபெறும் சிறு தொழில்கள், கைத்தொழில்கள் (household enterprises) பற்றிய விவரங்கள் சேகரிக்கப்பட்டன. எனவே, மொத்தம் 5 குடும்பங்கள் தெரிவு செய்ததுபோக, மீதம் எஞ்சி நிற்பவை 11 [16—5=11] குடும்பங்களாகும். அவற்றிலிருந்து, ஒரு (1) விவசாயத் தொடர்புடைய குடும்பமும், இரண்டு (2) விவசாயம் அல்லாத பணிகளில் ஈடுபடும் குடும்பங்களும் வாய்ப்பு அடிப்படையில் தேர்ந்தெடுக்கப்பட்டு, நுகர்வோர் (அல்லது துய்ப்பாளர்) செலவினங்கள் (consumer expenditure) பற்றிய விவரங்கள் பெறப்பட்டன. இதற்கேற்றபடி, ஒவ்வொரு நோக்கத்துக்கும் ஏற்றபடி விவரப்பட்டியல்கள் (schedules) தயாரிக்கப்படுவதை ஏற்கெனவே குறிப்பிட்டோம்.

இனி, கிராமியச் சூழ்நிலையில் (rural), மிகவும் முக்கியத்துவம் பெறும், நிலம் வைத்திருப்போர்பற்றிய ஆய்வினைக் குறித்து ஒருசிறு விளக்கத்தினைக் கருதுவோம்.

13. நிலம் பயன்படுத்தப்படும் விதம்பற்றிய அளவெடுப்பு (Land Utilization Survey)

கிராமப்புற நபர்கள் வைத்திருக்கும் நிலங்களைப் பயன்படுத்தும் விதங்களைப்பற்றிய விரிவான ஆய்வினை மேற்கொள்ள,

அந் நிலங்களை நேரடியாகக் கண்டு (observation) அதன்மூலம் விவரங்களைப் பெறுவதே சிறந்த வழிமுறையாகிறது. இதன் பொருட்டு ஒவ்வொரு கிராமத்திலும், ஒவ்வொன்றும் ஐந்து (5) நிலக் கூறுகள் (plots) கொண்ட 20 திரள்கள் (clusters), திரளின் பரப்பு (area) விகிதத்தில் தெரிவுசெய்யப்பட்டன. நிலக் கூறுகளின் எல்லைகளைக் (boundaries) காட்டும் கிராம வரைபடங்கள் இதற்குப் பயன்பட்டன. அதன்மூலம் நிலக் கூறுகள் முடிவு செய்யப்பட்டு, அதன் பின்னர்த் தேவையான அளவில் நிலக் கூறுகள் தேர்ந்தெடுக்கப்பட்டன. நிலத்தைப் பயன்படுத்தும் விதம் பற்றிய ஆய்வுக்கு, இவ்வாறு தெரிவுசெய்யப்பட்ட நிலப் பரப்புகள் எடுக்கப்பட்டன.

பயன்படுத்தப்படும் நிலம்பற்றிய ஆய்வுக்குத் தேவைப்படும் நிலக் கூறுகளின் திரள்களை, ஒழுங்குமாதிரி முறையினைப் பின்பற்றித் தேர்வு செய்யப்பட்ட கிராமங்கள் அனைத்திலிருந்தும் தெரிவுசெய்து, அதன்பின் இறுதிநிலைக் கூறுகள் எடுக்கப்படுகின்றன. கிராமங்களைத் தெரிவுசெய்கையில், மக்கள்தொகை அளவானது 1971-ல் நடைபெற்ற கணிப்பு (census) முடிவின் அடிப்படையில் கருதப்படுகின்றது. மதிப்பீடுகளில் தகுந்த மாற்றங்களும் ஏற்படுத்தப்படுகின்றன.

தெரிந்தெடுக்கப்பட்ட கிராமங்களை வரிசைப்படுத்திப் பட்டியல் அமைப்பதும், தகுந்தபடி ஈடுபடும் பணியின் (occupation) அடிப்படையில் படுகைகள் பிரித்து, சமவாய்ப்பு மாதிரி முறை மூலமாக வெவ்வேறு குடும்பங்களின் தன்மைகள்பற்றி மதிப்பிடுவதும் ஆய்வாளர்களின் பணியாக வகுக்கப்பட்டது. மக்களின் பொருளாதார, மற்றும் சமூகத் தன்மைகளைப் பகுத்து அறிதற்கு உதவி செய்யத்தக்க அளவில், அலுவல்கள் (அல்லது பணிகள்) அடிப்படையில் பிரிவினை செய்வது (occupational classification) அமைந்திருப்பது குறிப்பிடத்தக்கதாகும். நபர்களை அல்லது குடும்பங்களை மேற்கண்டவாறு தெரிவுசெய்தபின் பொருத்தமான பட்டியல்களைப் பயன்படுத்தி, நேரடித் தொடர்பு அல்லது நேரடி ஆய்வுமுறை மூலமாக விவரங்கள் திரட்டப்படுகின்றன. ஒவ்வொரு குடும்பத்திலும் ஒரு நபரை அல்லது சில நபர்களைத் தொடர்புகொண்டு, தேவையான விவரங்களை அறிந்து, அவ் விவரங்கள் அனைத்தும் ஆய்வாளர்களால் தமது பட்டியல்களில் குறித்துக்கொள்ளப்பட்டன.

NSS-ன் பொதுச் செயல்முறையானது (frame work) முதலிலேயே நாட்டின் பரப்பினை முழுதும் கருதி நடைமுறைப்படுத்தப்

பட்டது. தொடர்ந்துவரும் நிலைகளில், செயல்படும் விதமானது முன்னேறத்தக்க அளவில் முயற்சிகள் மேற்கொள்ளப்பட்டு வருகின்றன. பொதுவாக, நம் நாடானது பல்வேறு இனங்கள், மொழிகள், இயற்கை அமைப்புகள், வாழும் முறைகள், சமூக நிலைகள், பொருளாதாரத் தன்மைகள் ஆகியவற்றைக்கொண்டு அமைந்திருக்கும் சூழ்நிலையில், இத்தகைய விரிவானதோர் ஆய்வானது, எடுத்த எடுப்பிலேயே கருதப்படவேண்டியது இன்றியமையாததாகிறது. ஓரிடத்தில் பெறப்படும் அனுபவமானது இன்னோர் இடத்தில் பயன்தராது போகலாம். நடைமுறைச் சிக்கல்களும் இடத்துக்கு இடம் மாறுபடுவதைக் காணலாம். மேலும், இவ்வாறு பொது ஆய்வுமுறை மேற்கொள்ளப்படுவதால், காலதாமதம் ஏற்படுவதைத் தவிர்க்கலாம்; உடனடியாகச் சில பிரச்சினைகளைக் குறித்து முடிவெடுப்பது எளிதாகும். நாடு முழுதும் ஒரே சமயத்தில், விரிவான ஆய்வு மேற்கொள்ளப்படுவது நடைமுறை எளிமை கருதியே என்பதனைத் திட்டவட்டமாகக் கூறலாம்.

14. ஆய்வாளரின் பணியில் சில இடையூறுகள் (Difficulties of Enumeratcrs)

கள ஆய்வில் ஈடுபடும் காலத்தில், இடத்தின் சூழ்நிலைக்கு ஏற்ப, மற்றும் மக்களின் இயல்பிற்கு ஏற்ப, சில இடையூறுகளை ஆய்வாளர்கள் எதிர்நோக்க வேண்டியிருக்கும். அவற்றில் குறிப்பிட்ட சிலவற்றை இங்குக் காண்போம்:

(i) மக்களின் சில இயல்புகள்பற்றிய விவரங்களைச் சரிவர அறியமுடியாத சூழ்நிலை ஏற்படலாம். உதாரணமாக, ஒருவரது விருப்பு-வெறுப்புகள், பழக்கங்கள், சொந்த வாழ்வு-பற்றிய விவரங்கள், வரவு-செலவு விவரம் போன்றவைபற்றி எழுப்பப்படும் வினாக்களுக்குப் பதிலளிப்போர் உண்மையான பதில்களைத் (விவரங்களைத்) தராத நிலை ஏற்படும். எனவே, சாமர்த்தியமாக இவ் விவரங்களைப் பெறும்பொருட்டு, ஆய்வாளர்கட்குத் தேவையான, திறமையான பயிற்சி அளிக்கப்படுகிறது. இவ்வாறு சில இக்கட்டான நிலைகளில் சமயோசிதமாக விவரங்களைப் பெறுவதில் பெரும்பாலும் அவர்கள் வெற்றியும் பெறுகின்றனர். பெற்ற விவரங்கள் சரியா என்றறியத் தக்க வழிமுறைகள் இல்லை.

(ii) பொதுவாக, ஆய்வு நடத்தச் செல்லும் இடங்களில் வசிக்கும் மக்கள், ஆய்வாளர்களைச் சந்தேகக் கண்ணுடனே பார்க்கின்றனர். அதாவது, தங்களது செல்வநிலை, சொத்து விவரம்,

வருவாய் விவரம் போன்ற விவரங்களைப் பெறும்பொருட்டு, அரசினரால் அனுப்பப்பெறும் அலுவலர் என்றே அவர்களைக் கருதுகின்றனர். சொத்துகளைப் பறிமுதல் செய்ய அல்லது வருமான வரிவிதிக்க அரசு மேற்கொள்ளும் முன்னோடி நடவடிக்கை என எண்ணி ஐயுறுகின்றனர். எனவே, சரியான விவரங்கள் அதிகமாக வெளிவருவதில்லை. இதற்காக, அவர்களது முழு நம்பிக்கையைப் பெற்று விவரங்களைத் திரட்டுவதற்கு, ஆய்வாளர்கட்கு அதிக அளவில் திறனும்பொறுமையும் தேவைப்படுகின்றன. பொறுமை இழக்கநேரிடின், செயலானது முற்றுப்பெருமல் போகிறது. குறிப்பாக வயது, வருமானம், குழந்தைகள் எண்ணிக்கை போன்ற விவரங்களை முன்ஜாக்கிரதையுடன் (முன் எச்சரிக்கையுடன்) திரட்டுதல் வேண்டும்.

(iii) அடுத்த வகையான இன்னலானது, மொழிப் பிரச்சினை குறித்து அமைவதொன்ருகும். கிராமங்கள் அதிகமுள்ள நம் நாட்டில் இது எப்போதும் உள்ள பிரச்சினையாகும். ஆய்வு செய்யும் பகுதியில், வழக்கில் உள்ள மொழியினை (வட்டார மொழி) நன்கு அறிந்தவராக ஆய்வாளர் அமைந்திருப்பது, இந் நிலையில் உதவுகிறது. விவரங்கள் ஆங்கிலத்திலேயே குறிக்கப்பட்ட போதிலும், தேவையான விவரங்களைச் சேகரிக்கும் நிலையில், பதினைந்து மொழிகளுமே பயன்படுத்தப்பட்டன. அடிப்படை நோக்கமானது பாதிக்கப்படாது இருப்பதே குறிக்கோளாகக் கருதப்பட்டது.

(iv) இயல்புக்கு மாருன தட்ப வெப்ப நிலைகளாலும் (weather conditions), கொடிய விலங்குகள் வசிப்பதாலும் சில இடங்களுக்கு (கிராமங்களுக்கு) ஆய்வு நடத்தச் சென்று அடைவது இன்னல் மிகுந்த ஒன்ருகிறது. குறிப்பாக, இமாலயப் பகுதிகளில் சில இடங்களுக்குச் சென்றடைய, பனியானது உருகிப் பாதை சீர்படுவதற்கு நாள்கணக்கில் காத்திருக்கவேண்டிய நிலை ஏற்படுகிறது. சில இடங்களில், உயிருக்கே ஊறு நேரக்கூடிய சூழ்நிலையில், தகுந்த பாதுகாப்புடன் செல்வதும் அவசியமாகிறது. இது ஒரு பெரிய அளவினதான இன்னலாக அமைவது காணத்தக்கது.

(v) இவற்றேடல்லாமல், நாட்டின் வெவ்வேறு பகுதிகளில் காணப்படும் வெவ்வேறு பயிர்களின் பருவகால அளவுகளாலும் (crop seasons) இடர்ப்பாடுகள் நிகழ்கின்றன. ஆய்வானது ஒரே சீரானதாகஅமையும்படிநேரிடாமல் போகின்றது. பயிர்ச் சாகுபடி பற்றிய ஆய்வுகள் (crop surveys) பெரிதும் முக்கியத்துவம் பெறு

வதால், இதனைக் கருதுவதும் இன்றியமையாததாகும். மேலும், இடத்திற்கு ஏற்றபடி மாறுபடும் எடைகள் (weights) மற்றும் அளவீடுகள் இவற்றின்மூலமும், பொதுவாகக் கிராமங்களைப் பொறுத்த அளவில், சில மதிப்பீடுகளைக் கருதுகையில் பிரச்சினைகள் எழுகின்றன.

எல்லாவித இடர்ப்பாடுகளையும் திறமையுடன் சமாளித்து, தேவையான விவரங்களனைத்தையும் தடையின்றிப் பெறுதற்குப் பெரிதும் முயற்சிகள் எடுக்கப்பட்டு, NSS வழிமுறைகள் செம்மையாக நிறைவேற்றப்பட்டு வருகின்றன.

15. NSSபற்றிச் சில முக்கியக் குறிப்புகள்

கிராமப்புறங்களின் தன்மைமட்டும் அன்றி, நகர்ப்புறங்களையும் விரிவான அளவில் எடுத்துக்கொண்டு ஆய்வுகள் மேற்கொள்ளப்படும் தன்மையதான NSS ஆனது, நம் நாட்டின் பல்வேறு துறைகளின் வளர்ச்சிக்கு இன்றியமையாததாக ஆகிறது. இன்றுவரை 27 சுற்றுகளை நடத்தி முடித்திருக்கும் NSS இயக்கச் செயலகமானது, (NSS Directorate) தொடக்கத்திலிருந்தே சிறந்த பணியாற்றி வந்திருக்கின்றது. மக்களது பொருளாதார, சமூக, நிதி போன்ற இயல்புகளின் முழு விவரங்களும், பிறப்பு-இறப்புபற்றிய விவரங்களும்,இன்னும் பல முக்கியப் புள்ளி விவரங்களும் சேகரிக்கப்பட்டுப் பல நோக்கங்களுக்குப் பயன் படுத்தப்படுவதால்,இன்னும் அதிக அளவில் முழுக் கவனத்துடன் இவ்வியக்கம் பணியாற்ற வேண்டியது அவசியமாகிறது.

பல்வேறு இன்னல்களின் இடையே திறம்படப் பணியாற்றும் NSS அமைப்பின் செயல்முறை பாராட்டத்தக்கதொன்றாகும். இனிவரும் சுற்றுகளிலும் இதன் பணியானது மிகவும் செம்மையுடன் விளங்கும் என எதிர்பார்க்கப்படுகிறது. மிக முக்கியமாக ஆய்வுக்குக் கருதப்படும் காலவரைகள் (period of survey) ஒவ்வொரு சுற்றிலும் அடிக்கடி மாற்றப்படுவதால், பல்வேறு முடிவுகளை ஒப்பிட்டுப் பார்ப்பது இயலுவதில்லை. எனவே, கூடுமானவரையில் ஒரே மாதிரியான ஆய்வுக் காலவரையினை இனிவரும் சுற்றுகளில் கடைப்பிடிப்பது, சில பிரச்சினைகளைத் தவிர்க்கும் என எதிர்பார்க்கப்படுகிறது.

கிராமப்புறங்களில் (ஏறக்குறைய 5,60,000 கிராமங்கள் அமைந்துள்ளன) மக்களது ஒத்துழைப்புப் பல்வேறு காரணங்களால் சரியான அளவில் கிடைப்பதில்லை. ஆதலால், NSS ஆய்வுச் செயல்முறைகளிலும், கள ஆய்வுகளிலும் காலதாமதம்

ஏற்படுகிறது. எனவே, இதனைத் தவிர்க்க, வழிமுறைகளை ஆராய்ந்து கண்டறிந்து, NSS-ன் செயலாற்றும் வேகம் அதிகரிக்கப்படின், அதிகப் பயனளிக்கும் எனலாம். அதன் நோக்கமும் இன்னும் பரந்த அளவிற்கு விரிவாக்கப்படின், சிறந்த முடிவுகள் கிடைக்கும் என்பதும் கருத்திற்கொள்ளப்படவேண்டியதாகும். மேலும், விவரங்களைப் பெறப் பயன்படுத்தப்படும் வினாப் பட்டியல்கள் (Questionnaires or Schedules) எளிதில் புரிந்து கொள்ளத்தக்க விதத்திலும், சிக்கல் ஏதும் இல்லாமலும் அமைக்கப்படவேண்டியது மிக இன்றியமையாததாகும். அதன்மூலம், பதிலளிப்போர் தாமாகவே முன்வந்து விவரங்களைத் தரும் அளவிற்குக் கணிப்பு எளிதாகும் என்பது காணத்தக்கது.

எனவே, ஒருசில நடவடிக்கைகளைச் சிறிது கவனத்துடன் எடுப்பதன்மூலம் NSS ஆய்வுப் பணி செம்மைப்படுத்தப்பட்டு, அதன் முடிவுகள் அனைத்தும் நாட்டின் பல துறைகளுக்கும் பெரிதும் பயன்படும் என்பதில் ஐயமில்லை.

II. மக்கள்தொகைக் கணிப்பு ஆய்வு (POPULATION CENSUS)

[நோக்கம்—இரு வழிமுறைகள்—இந்திய மக்கட் கணிப்பு —செயல் முறை அமைப்பு—1951-ல் மக்கட் கணிப்பு—1961-ல் மக்கட் கணிப்பு —1961கணிப்பில் சில இயல்புகள்—1971 மக்கட் கணிப்புஆய்வு—முக்கிய அம்சங்கள்—தனிநபர் ஆய்வுச் சீட்டு—1971-ல் பெறப்பட்ட சில விவரங்கள்—நடைமுறை இடர்ப்பாடுகள்—மக்கட் கணிப்பின் தேவையும், பயனும்.]

1. மக்கள்தொகைக் கணிப்பின் நோக்கம் (Need)

சமூக மற்றும் பொருளாதார இயல்புகளைப் பொறுத்து மனிதரது வளர்ச்சியானது, எந்த ஒரு நாட்டிற்கும் அதன் மொத்த மக்கள்தொகையானது வளர்வதையோ அல்லது தேய்வதையோ பொறுத்து அமையத்தக்கதாகும். எனவே. மக்கள்தொகைக் கணக்கெடுப்பானது (census of population) அதிக அளவில் கவனம் செலுத்தப்பட்டு நடத்தப்படுகிறது. ஒரு குறித்த கால அளவில், சமூக மக்களின் பல்வேறு இயல்புகளைப்பற்றி நன்கு படம்பிடித்துக்காட்ட இது மிக இன்றியமையாததாகும்.

பொதுவாக, மக்கள்தொகைக் கணிப்பினை, 'ஒரு குறிப்பிட்ட காலத்தில், ஒரு நாட்டின் அனைத்து மக்களையும் கருதி, அவர்தம்

சமூக, பொருளாதார, பிறப்பு-இறப்புப் போன்ற எல்லா இயல்பு களையும்பற்றிய விவரங்களைத் திரட்டி, அவற்றைத் தக்கபடி தொகுத்து, இறுதி முடிவுகளை வெளியிடும்' ஓர் ஆய்வாக வரை யறுக்கலாம். இது மிக இன்றியமையாததொரு முழுக் கணிப்பு ஆய்வுமுறை (complete enumeration) ஆகும். இதன் நோக்கம் (scope) மிகப் பரந்த ஒன்று ; மக்கள்தொகை அளவினைமட்டும் கணக்கிடுவது இதன் அடிப்படை நோக்கமன்று ; மக்களைப் பற்றிய வெவ்வேறு வகையான விவரங்களை ஒன்றுவிடாமல் சேகரித்து, பல்வேறு துறைகளில் பயன்படுத்துவதே முக்கிய நோக்கமாக அமைகிறது. சமுதாய வளர்ச்சி, பொருளாதார முன்னேற்றம், மக்கள் நலன், தொழில் அபிவிருத்தி, வெளிநாட்டு வர்த்தக முன்னேற்றம் போன்ற பல்வேறு பிரச்சினைகளும் இதனுடன் நெருக்கமாகப் பிணைக்கப்பட்டிருக்கின்றன. எதிர் காலத்துக்குத் தேவையான அளவில் திட்டங்கள் வகுப்பதும், விதிமுறைகளை உருவாக்குவதும் இதன் அடிப்படையில் அமைவ தால், தனியான அளவில் நிறையக் கவனம் செலுத்தப்பட்டு ஆய்வுகள் மேற்கொள்ளப்படுகின்றன.

2. மக்கள்தொகைக் கணிப்பின் இரு வழிமுறைகள் (Two Methods)

பொதுவாக, மக்கள்தொகைக் கணக்கெடுப்பில் இருவகையான முறைகள் கையாளப்படுகின்றன. அவற்றைச் சுருக்கமாக இங்குக் காண்போம்:

(i) டி ஃபாக்டோ முறை (De-facto Method) : இந்தமுறையின் படி, கணக்கெடுப்பு நடத்தப்படும் நாள் அல்லது இரவு நேரத்தில் ஒருவர் எந்த இடத்தில் இருக்கிறாரோ, அந்த இடத்திலேயே கணக் கெடுக்கப்படுகிறார். இது 'நாள்பற்றிய முறை' (Data System) என்று குறிப்பிடப்படும். இந்த முறையில், ஒரே இரவில் நாடு முழுதும் ஒரே சமயத்தில் கருதப்பட்டுக் கணிப்பு நடைபெற வேண்டும். இது நடைமுறையில் எளிதாகவும், நாட்டில் வசிக்கும் (குறிப்பிட்ட நாளில்) மக்கள் பற்றிய மதிப்பீடுகள் பெறும்படியாக வும் அமைகிறது. நாடுகளின் இடையில் மேற்கொள்ளப்படும் ஒப்பிடுதலும் (comparison) இதன்மூலம் எளிதாகிறது. நாடு முழுவதும் ஒரே சமயத்தில் கணக்கெடுப்பு நடத்தப் படுவதால் இது இயலுகிறது.

ஆனால் இந்த வழிமுறையில் குறைபாடுகள் (drawbacks) அதிக அளவில் இடம்பெற நேரிடும். தாற்காலிகமாக ஒரு

நாட்டில் வசிக்கும் குடிமக்கள் கணக்கீட்டில் இடம் பெறவும், அந்தக் காலத்தில் நாட்டைவிட்டு வெளியேறி இருக்கும் குடிமக்கள் கணக்கெடுக்கப்படாமலும் இருக்கும் நிலை ஏற்படுகிறது. எனவே, ஒரு நாட்டின் நிரந்தரமான மக்கள் தொகையினை முடிவு செய்ய இயலுவதில்லை. வெவ்வேறு பகுதிகட்கு (regions) ஏற்ற அளவில் பிரித்து ஆய்வுமுறை பயன் படுத்த முடிவதில்லை. மேலும், கணக்கெடுப்புக் காலத்தில் பயணத்தில் ஈடுபட்டிருப்பவர்களையும் வனப் பிரதேசங்களில் வசிப்பவர்களையும் இராணுவப் பணியில் ஈடுபட்டிருப்பவர்களையும் கணக்கெடுக்க இயலாத சூழ்நிலை அமைகிறது. குறித்த காலத்தில், குறுகியகாலவரையில் ஆய்வுகளை நடத்திமுடிக்கப் பெரும் அளவில் முயற்சி தேவைப்படும். எனவே, நடைமுறையில் இம்முறை அதிகமாகப் பின்பற்றப்படுவதில்லை.

(ii) டி ஜூரி முறை (De jure Method) : இந்த முறையானது, முதல் முறைக்குமுற்றிலும் எதிரானது. ஒரு குறிப்பீட்டுக்காலத்தில் (reference period) ஒரு நபர் இயல்பாக வசிப்பதாகக் கூறப்படும் அல்லது அறியப்படும் இருப்பிடத்தில் (household) அவரைப் பற்றிக் கணக்கெடுப்பு நடத்தப்படும். அந்த இடத்தில் ஆய்வு நாளன்று அவர் இருந்தாலும் இல்லாவிட்டாலும் அவரைப்பற்றிக் கணக்கெடுப்புநடத்தப்படும். எனவே, ஒரு நபரைப்பற்றிய கணக் கெடுப்பு அவர் பணியாற்றும் அலுவலகத்திலோ, நிறுவனத்திலோ, பயிலும் கல்விக்கூடத்திலோ, புகைவண்டி நிலையத்திலோ, அல்லது வேறு இடங்களிலோ நடத்தப்படுவதில்லை. ஒவ்வொரு குடும்பத்தினைப் பற்றியும் அந்தந்த இருப்பிடங்களைக் கருதி, கணக்கெடுப்பு நடைபெறுகிறது. பல நபர்கள் ஒன்றாக வசிக்கும் படியும், ஒரே இடத்தில் உணவு உட்கொள்ளும் நிலையிலும் அமையும் தன்மையானது இயல்பாக ஒரு 'குடும்பம்' என்று கருதப்படலாம்.

இம் முறையினைக் 'காலவரை பற்றிய கணக்கெடுப்பு முறை' (Period System) என்றும் கொள்ளலாம். ஏனெனில், ஓரிரவில் நடத்தி முடிக்கும் தன்மையதாக இல்லாமல், ஒரு குறிப்பிட்ட கால அளவில் (period) பரந்து அமையத்தக்கவிதத்தில் கணிப்பு நடத்தப் பெறுவதால், இவ்வாறு குறிப்பிடலாம். முன் குறிப்பிட்ட முறையினைவிட, இதன்மூலம் மிகவும் அதிகத் துல்லியமான முடிவுகளைப் பெற இயலுகிறது. ஒரு குறிப்பிட்ட நிலப்பரப்பில் ஒவ்வொரு கணக்கெடுப்பாளரும் ஆய்வுப்பணி மேற்கொள்வதால், பெரிய அளவில் விரிவான பரப்பில் கணிப்பு

ஆய்வு நடத்திமுடிக்கப்பட வழி ஏற்படுகிறது. விவரங்கள் விட்டுப்போக நேரிடாமலும் (omission) மற்றும் ஒரே விவரமானது திரும்பவும் சேகரிக்கப்படாமலும் (duplication) இருக்கும் நிலை காணப்படும். பூகோளப் பரப்பின் அடிப்படையிலும், வட்டார (region) அடிப்படையிலும் மக்கள் வெவ்வேறு பகுதிகளாகப் பிரிக்கப்பட்டு, சமூக, அரசியல் மற்றும் பொருளாதாரப் பிரச்சினைகளை நன்கு ஆய்ந்து கண்டறிய இம்முறையானது பெரிதும் உதவுகிறது. மேலும், வெவ்வேறு பகுதிகட்கும் (regions) நிதி ஒதுக்குவது, வரி விதிப்பு நிர்ணயிப்பது போன்ற பல பணிகளும் எளிதாகின்றன.

இம்முறையிலும் சில குறைபாடுகள் காணப்படும். நிரந்தர இருப்பிடம் இல்லாதவர்கள், அங்குமிங்கும் அலைவோர் போன்ற பலரைக் கணக்கிடுவது சற்றுச் சிரமமான காரியமாக அமைகிறது. ஒன்றுக்கு மேற்பட்ட இருப்பிடங்களைக் கொண்டிருப்போர் மூலமும் பிரச்சினை ஏற்படும். சில வரையறைகள் எளிதில் புரிந்துகொள்ளமுடியாதபடி வகுக்கப்பட நேரிடலாம். இந்திய நாட்டு மக்கள்தொகைக் கணக்கெடுப்புகளில் இம் முறையே முழுதும் பின்பற்றப்பட்டு வருகின்றது.

கணக்கெடுக்கும் விதம்பற்றிய ஒரு சிறுகுறிப்பு

இருவழிகளில் மக்களிடம் விவரங்கள் திரட்டப்படுகின்றன. முதலாவது, தனிநபர் (குடும்பநபர்) கணக்கெடுப்பு என்பதாகவும் இரண்டாவது ஆய்வாளர் கணக்கெடுப்பு (Self enumeration and canvasser enumeration) என்பதாகவும் அமைகின்றன.

முதல் வழியில், குடும்பத் தலைவரிடம் வினாப்பட்டியல் தரப்பட்டு, விவரங்கள் அவர்மூலம் குறிக்கப்பட்டு, ஆய்வாளர்களால் திரும்பப் பெறப்படுகின்றன. கல்வி அறிவு அதிக அளவில் காணப்படும் பகுதிகளில், இம்முறை பயன்தரும்.

இரண்டாவது வழியில், நிறைய அளவில் கணக்கெடுப்போர்கள் (enumerators) தேவைப்படுகின்றனர். ஏனெனில், ஒவ்வொரு குடும்ப நபரையும் ஆய்வாளர் அணுகி, நேரடித் தொடர்புகொண்டு, அவர்கள்மூலம் அளிக்கப்படும் விவரங்களைத் தாமே பட்டியலில் குறித்துக்கொள்ள வேண்டியிருக்கிறது. குறிப்பாக, கல்வி அளவு குறைந்து காணப்படும் கிராமப்புறங்களில் இம்முறை மிகவும் தேவைப்படும் ஒன்றாகிறது. தகுந்த பயிற்சி பெற்ற ஆய்வாளர்களால், தெளிவான விவரங்களைப் பெற இம் முறையானது வழிசெய்கிறது. ஆனால், பெரும் அளவில் ஆய்வுச் செயல்முறை திட்டமிட வேண்டியிருப்ப

தால் இத்தகைய வழிமுறையானது மிகுந்த செலவு உள்ளதாக அமைகிறது. விரிவான முறையில் திட்டங்களை அமைத்துக் கொண்டு செயல்படவேண்டியிருக்கும். தவிர்க்கமுடியாத இடங்களில் இம்முறையினைப் பின்பற்றுவதுசிறந்ததாகும். வீடுகள் அனைத்தும் ஓர் ஒழுங்கின் அடிப்படையில் இலக்கமிடப்படுவது (numbering) கணக்கெடுப்பின் முதற் கட்டப் பணியாகும்.

இந்திய மக்கள்தொகைக் கணிப்பு (Indian Population Census)

இந்திய நாட்டின் மக்கள்தொகைக் கணிப்பானது பத்து ஆண்டுகளுக்கு ஒருமுறை (decennial) மேற்கொள்ளப்படுவ தாகும். மிகவும் விரிவான அளவில் மேற்கொள்ளப்படும் முழு ஆய்வுகளில் மிகச் சிறந்ததாக இது விளங்குகிறது. அண்மையில் நடைபெற்ற மக்கள்தொகைக் கணிப்பானது (1971), வரிசையில் அமையும் பத்தாவது கணக்கெடுப்பு ஆகும். முதல் கணக்கெடுப்பு 1872-ல் நடத்தப்பட்டது. எனவே மக்கள் தொகைக் கணிப்பின் ஒரு நூற்றுண்டுக் காலப் (centenary) பூர்த்தியினைக் காணலாம்.

1948ஆம்ஆண்டில், 'இந்தியமக்கள்தொகைக்கணக்கெடுப்புச் சட்டம்' (Indian Census Act) உருவாக்கப்பட்டு நிரந்தரமாக அமைக்கப்பட்டது. இதன்மூலம், தொடர்ந்து கணிப்புகள் மேற்கொள்ளப்படவும் நோக்கங்கள் விரிவாக்கப்படவும் வழி ஏற்படுத்தப்பட்டது. இந்தச் சட்டமானது, வெவ்வேறு மட்டங் களில் ஆய்வுகட்குத் தேவைப்படும் அலுவலர்களையும், அதிகாரி களையும் நியமிக்க மத்திய அரசுக்கு முழு உரிமையினை அளிப்ப தாகும். மேலும் வெவ்வேறு பணிகளின் செயல்முறைகள் இதன் மூலம் உருவாக்கப்படுகின்றன. இச் சட்டத்தின்படி, ஒவ்வொரு குடிமகனும் தேவையான விவரங்களை உண்மையுடன் அளிக்கக் கடமையுடையவனுகை ஆகிருன். உண்மை அனைத்தும் வெளி யிடப்படாது இரகசியமாக வைக்கப்படுகின்றன. ஆய்வாளர்கள், ஒவ்வொரு வீட்டிற்கும்சென்று தேவையான விவரங்களைப் பெறவும் உரிமை அளிக்கப்படுகின்றனர். எந்தவித விவரமும் தனியான (individual) முக்கியத்துவம் பெருது இறுதிநிலைப் பட்டியல்கள் அமைக்கப்பட உதவும் கூறுகவே கருதப்படும். கணிப்பு ஆய்வினை வெற்றிகரமாகத் திட்டமிட்டு முடிக்க இந்தச் சட்ட மானது மிகவும் உதவிபுரிகிறது.

1. செயல்முறை அமைப்பு (Organisation)

இந்திய அரசானது (Govt. of India), மக்கள்தொகைக் கணக்கெடுப்பின் மேல்மட்டத்தில் 'பொதுப் பதிவாளராகவும்',

'கணக்கெடுப்பு உயர்மட்ட அதிகாரியாகவும்' (ஆணையர்) (Registrar General and Ex-officio, Census Commissioner) அமைந்திருக்கும் ஒரு மேல்மட்ட அதிகாரியை நியமனம் செய்கிறது. இந்த அதிகாரியின்கீழ் முழு ஆய்வும் நடைபெறுகிறது. இவர் மூலம், மக்கட் கணிப்பின் செயல்முறைகள் அனைத்தும் விரிவான அளவில் திட்டமிடப்பட்டு, வெவ்வேறு மத்திய, மற்றும் மாநில அரசு அலுவலர்களைக் கொண்டு ஆய்வு நடத்தப்படுகின்றது.

ஒவ்வொரு மாநிலத்திலும் ஆய்வுப் பணியினை நடத்தக் கணிப்பாய்வு இயக்குநர்கள் (Directors of Census) நியமிக்கப் படுகின்றனர். ஆய்வுப் பணிகளை வழிநடத்துவது (directing) இவர்தம் கடமையாகும். மாநில அரசின் துணையுடன், மாநில அரசில் பணியாற்றும் பல்வேறு அலுவலர்கள்மூலம் ஒவ்வொரு மாநில அளவில் ஆய்வுப் பணிகள் கருதப்படுகின்றன.

அடுத்து, ஒவ்வொரு மாவட்டத்திலும் நடைபெறும் மக்கள் தொகைக் கணிப்பிற்கு, அம் மாவட்டத்தின் ஆட்சித் தலைவர் (District Collector) பொறுப்பேற்கிறார். இதன் பின்னர், ஒவ்வொரு மாவட்டத்திலும் அடங்கும் வட்டப்பிரிவு (circle), தாலுகா, நகராட்சி (municipality) போன்றவை தனித்தனியாக ஆய்வுப் பணிக்குப் பொறுப்பேற்கின்றன. தாசில்தார், நகராட்சி ஆணையர் (Municipal Commissioner) போன்றோரின் தலைமையில், இம்மட்ட ஆய்வுப் பணிகள் நிறைவேற்றப்படும்.

அடுத்த கட்டமாக, மேற்குறிப்பிடப்பட்ட ஒவ்வொரு பொறுப்புக் களமும் (charge), நான்கு முதல் ஐந்துவரை (4—5) ஆய்வாளர் தொகுதிகளைக் (blocks) கொண்ட பல உள் வட்டங்களாகப் (circles) பிரிக்கப்பட்டு, ஒவ்வோர் உள்வட்டத்திலும் ஒரு மேற்பார்வையாளர் (supervisor) நியமிக்கப்படுவார். கருதப்படும் 'ஆய்வாளர் தொகுதி'யானது, (enumerator's block) 100-லிருந்து 150 வீடுகளைக் கொண்டதாக, 750 முதல் 1000 வரையிலான (சுமார்) மக்கள்தொகை உள்ளதாக அமைவது குறிப்பிடத்தக்கது. இத்தகைய எல்லை நிர்ணயிக்கப் பட இயலாவிடின், தொகுதியினை ஒரு தனி ஆய்வாளரால் ஆய்வு மேற்கொள்ளத் தக்கவிதத்தில் அடக்கமான அளவில் அமைத்தல் கவனிக்கத்தக்கது. குறிப்பாக, இத் தொகுதி ஏதேனும் கிராமம் அல்லது பேரூர் (Panchayat) எல்லையைக் கடக்காததாக அமையவேண்டும்.

மேலும், குறிப்பான சில இயல்புகளில், அதாவது பெரிய தொழில் நிறுவனங்கள், குடியிருப்புகள் (colonies) இராணுவம், தங்கும் விடுதிகள் போன்றவற்றினைக் கருதும் நிலைகளில், சில 'பொறுப்பு'களை (charge) உருவாக்குதல் தேவைப்படும். இவற்றிற்கு மிகவும் பொறுப்புவாய்ந்த அலுவலர்கள் நியமிக்கப் படுகின்றனர். இவற்றின் பணிகள் அனைத்தும், மாவட்டக் கணிப்பு அதிகாரிகளின் கட்டுப்பாட்டிற்கு உட்பட்டனவாகும். கணக்கெடுப்பு ஆய்வாளர்கள், கிராம அலுவலர்கள், ஆசிரியர், அரசு ஊழியர் போன்ற பலரிடையே தேர்ந்தெடுக்கப்பட்டு, ஆய்வுப் பணிகள் நடத்தப்படுகின்றன. இப்பணியாளர்கட்குச் சிறு அளவில் ஊதியங்கள் (honorarium) அளிக்கப்படுகின்றன. அதாவது, இத்தகைய மக்கட் கணிப்பு ஆய்வுப் பணிகள் அனைத்துமே சமூகப் பணிகளாகக் (social services) கருதப் படுகின்றன. இவர்கட்குத் தேவையான பயிற்சிகளும் விவரக் குறிப்புகளும் (instructions) தீவிரமான அளவில் அளிக்கப் படுகின்றன.

2. 1951-ல் நடைபெற்ற மக்கட் கணிப்பு (Census, 1951)

சுதந்திர இந்தியாவின் முதலாவது மக்கட் கணிப்பு 1951-ல் நடைபெற்றது. ஜம்மு-காஷ்மீர்ப் பகுதி தவிர, நாட்டின் ஏனைய பகுதிகள் அனைத்தும் கருதப்பட்டு விரிவாய்வு மேற்கொள்ளப் பட்டன. கணக்கெடுப்போரால் பயன்படுத்தப்பட்ட 'ஆய்வுச் சீட்டுகளில்' (enumeration slips) விவரங்கள் குறிக்கப்பட்டு, ஒவ்வொரு நபர்பற்றியும் பின்வரும் விவரங்கள் பெறப்பட்டன :

1. பெயரும், குடும்பத் தலைவருடன் உறவு முறையும் :
2. (அ) எந்த நாட்டைச் சேர்ந்தவர் :
   (ஆ) மதம் :
   (இ) மத உட்பிரிவு :
3. வாழும் நிலை :
4. வயது :
5. பிறந்த ஊர் :
6. வெளியேறி வந்தவராக இருப்பின், வந்து சேர்ந்த தேதி:
   பாகிஸ்தானில் இருந்த மாவட்ட விவரம் :
7. தாய்மொழி :
8. இரு மொழி அறிந்தவரா? :

9. பணி புரிபவரா? அண்டியிருப்பவரா?
(Dependency) :

10. வாழ்விற்கு வழிவகுக்கும் அடிப்படைக் கூறு :
(Principal means)

11. வாழ்விற்கு வழிவகுக்கும் துணைக்கூறு :
(Secondary means)

12. எழுதப்படிக்கத் தெரிந்தவரா ?
மற்றும் கல்வியறிவு நிலை :

13. வேலையற்றவரா (Unemployment) :

14. இனம் (Sex) :

இவற்றில் 13ஆவது கேள்வியைத் தவிர, ஏனைய வினுக்கள், அனைத்து மாநிலங்கட்கும் பொதுவானவை. 13ஆவது வினுவினை, ஏதேனும் பொருள்பற்றி விவரம் பெறத்தக்க விதத்தில் ஒவ்வொரு மாநிலமும் தன் இச்சைப்படி அமைத்துக் கொள்ள வழி செய்யப்பட்டது. சென்னை, பம்பாய் மாநிலங்களில், வளமை (fertility) பற்றிய விவரம் பெறப்பட்டது.

இந்த ஆய்வில், ஆய்வுக் காலவரை (enumeration period) நிர்ணயிக்கப்பட்டது. மூன்றுவாரக் கால அளவினதாக அமைக்கப்பட்டு, அதன் பின்னர், ஒரு வாரம் மறுபரிசீலனை, மற்றும் சரிபார்த்தல் ஆகியவற்றிற்குக் கருதப்பட்டது. மேலும், சில வரையறைகள் உருவாக்கப்பட்டன. அதன்படி ‘ வீடு ’ என்பது ஒரு தனியான முக்கிய நுழைவாயில் கொண்ட வாழும் இடமாகக் (dwelling place) கருதப்பட்டது ; ‘குடும்பம்’ என்பது ஒன்றுக ஒரே இடத்தில் வசிப்பதாகவும், ஒரு பொதுவான ‘ சமையலறை’யினின்று பெறப்படும் உணவினை உட்கொள்ளும் படியாகவும் அமையும் தனி நபர்களை அல்லது நபர்கள் கொண்ட குழுவினைக் குறிப்பதாக வரைமுறை செய்யப்பட்டது. இதன்படி, ஆய்வு நடத்தப்பட்டது. மற்றும், தேசிய அளவில் பதிவேடு அமைக்கப்பட்டது.

3. 1961-ல் நடத்தப்பட்ட இந்திய மக்கட் கணிப்பு (Indian Census-1961)

இதில் பயன்படுத்தப்பட்ட தனி நபர் ஆய்வுச் சீட்டானது, பின்வருமாறு விரிவான முறையில் அமையப்பெற்றது :

இந்திய மக்கட் கணிப்பு—1961

(இரகசியமானது)

இடத்தின் குறியீடு
(Location Code)

அட்டை எண்:
சீட்டு எண் :
குடும்ப எண் :

1. (அ) பெயர் (முழுமையாக)...
(ஆ) குடும்பத் தலைவருடன் உறவு முறை...

2. வயது ... (சென்ற பிறந்த நாளின் போது)

3. மணமானவரா...

4. (அ) பிறந்த ஊர்...
(ஆ) பிறந்தது (Born) .. கிராமப்புறம்/நகர்புறம்...

4. (இ) வேறு இடத்தில் பிறந்திருப்பின தங்கியிருந்த காலம் (Duration)

5. (அ) எந்த நாட்டைச் சேர்ந்தவர்... (Nationality) பிற்பட்ட வகுப்பினரா (S. C.)... தாழ்த்தப்பட்டவரா (S.T.)...

5. (ஆ) மதம்...

6. எழுதப்படிக்கத் தெரிந்தவரா ? கல்வி நிலை...

7. (அ) தாய்மொழி...
(ஆ) வேறு ஏதேனும் அறிந்த மொழி...

8. பயிரிடுபவரா... (Working as cultivator)

9. விவசாயப் பணியாளரா... (Agricultural labourer)
(அ) பணியின் தன்மை... (nature)
(ஆ) குடும்பத்தில் மேற்கொள்ளப் படும் சிறுதொழில்... (Household industry)

(இ) வேறு பணியாளராக இருப்பின்
(i) வேலையின்தன்மை...
(ii) தொழில், வர்த்தகம் பணி பற்றிய விவரம்...
(iii) கூலி பெறுபவராக இருப்பின், எந்தவகை (Class)...
(iv) நிறுவனத்தின் தன்மை... (Nature of establishment)

10. குடும்பங்களில் நடைபெறும் தொழில்களில் பணி புரிபவரா... (Household industries)
(அ) வேலையின் தன்மை
(ஆ) தொழில், பணி ஆகியவைபற்றி...

11. (8), (9), (10) இவற்றைத்தவிர இதர பணிகள் புரிபவரா...

12. வேலையற்ற நிலையில், ஈடுபடும் செயல்...

13. இனம் (sex) ..

4. 1961 கணிப்பின் சில குறிப்பிட்ட சிறப்பு இயல்புகள் (Special Features)

1961 கணக்கெடுப்பானது, இந்தியப் பொருளாதார மூன்று வது ஐந்தாண்டுத் திட்டத்தின் (Third Five-Year Plan) தொடக்கத்தை ஒட்டியதாக அமைந்தது. இது ஒரு குறிப்பிடத்தக்க ஆய்வாக அமைகிறது ; ஏனெனில், முதல் இரண்டு ஐந்தாண்டுத் திட்டங்களிலும் கருதப்பட்ட பல்வேறு ஆக்கப்பணிகளின் பயன்களை (gains) நன்கு மதிப்பிட வசதி ஏற்பட்டது. இயல்பான காலத்தில், நாட்டின் அமைதியான சூழ்நிலையில், விரிவாகவும் முழுமையாகவும் நடைபெற்ற முதல் ஆய்வாக இது விளங்குகிறது; பின்வரும் சிறப்பியல்புகள் கொண்டதாக அமைகிறது :

(1) 'வீடு' மற்றும் 'குடும்பம்' பற்றிய வரையறைகள் மேலும் விரிவாக்கப்பட்டன. 'வீடு' என்ற கருத்தானது, 'கட்டடம் எனவும்', 'கணக்கெடுப்பு இல்லம்' எனவும் (building and census house) இரண்டாகப் பிரிக்கப்பட்டது. நிலப்பரப்பில் (ground) தரையினில் எழுப்பப்படும் முழு அமைப்பானது (structure) 'கட்டடம்' எனவும், தனியான ஒரு நுழை வாயிலைக் கொண்டு அமையும் கட்டடம் அல்லது கட்டடத்தின் ஒரு பகுதியானது, 'கணக்கெடுப்பு இல்லம்' என்பதாகவும் வரைமுறைகள் செய்யப்பட்டன. எனவே, ஒரு கட்டடத்தில் பல 'கணக்கெடுப்பு இல்லங்கள்' இருக்கும் நிலையைக் காணலாம். 'குடும்பம்' பற்றிய வரையறை முன்புபோலவே கருதப்பட்டது.

(2) குடும்பத்தில் நடைபெறும் தொழில் (அல்லது வீட்டுத் தொழில்) பற்றிய வினவானது முதன்முறையாக 1961 கணிப்பில் இடம்பெற்றது. குடும்பம்பற்றிய பொருளாதார விவரங்கள், புதிதாக உருவாக்கப்பட்ட 'குடும்பப் பட்டியல்' (household schedule) மூலமாகப் பெறப்பட்டன. இத்தகைய விவரப் பட்டியலானது குடும்பங்களில் மேற்கொள்ளப்படும், அதாவது குடும்ப நபர்கள் மேற்கொள்ளும் சிறு தொழில்கள் (small scale industries), பயிர்த் தொழில் (cultivation) போன்றவை பற்றிய விவரங்களைப் பெறத்தக்க விதத்தில் அமைக்கப்பட்டது.

(3) 'வீடுகளின் வரிசைப்பட்டியல்' (house list) தயாரிக்கப்பட்டது, இக் கணிப்பின் மற்றொரு சிறப்பியல்பாக அமைகிறது. ஒரு தரமான படிவம் தயாரிக்கப்பட்டு (standard form), அதன்மூலமாகக் 'கணக்கெடுப்பு இல்லம்' பயன்படுத்தப்படும் விதம், தள அமைப்பிலும் கூரை அமைப்பிலும் (flooring and

roofing) பயன்படுத்தப்பட்ட பொருள்கள், சொந்த வீடா குடியிருக்கும் வீடா என்னும் தெளிவு, வசிக்கும் அறைகள் என்பனபோன்ற பல்வேறு விவரங்களும் சேகரிக்கப்பட்டன. இதோடுகூட நிறுவனங்கள் (establishments), தொழிற்கூடங்கள் (factories) போன்றவற்றிற்குப் பயன்படுத்தப்பட்ட 'இல்லங்களை'ப் (houses) பொறுத்த அளவில் உரிமையாளர் அல்லது நிறுவியோர், தயாரிக்கப்படும் பொருள் மற்றும் உற்பத்தி அளவு, உழைப்பாளர் எண்ணிக்கை, பயன்படும் மூலப்பொருள் போன்ற பல்வேறு தரப்பட்ட விவரங்களும் முழு அளவில் திரட்டப்பட்டன.

(4) மக்களின் பொருளாதார இயல்புகட்கு அதிக முக்கியத்துவம் அளிக்கப்பட்டது. இந்த அடிப்படையில் அமைந்த பாகுபாடானது (classification) மிகவும் மாற்றப்பட்டது. 1951-ல் (i) வேலை செய்வோர், (ii) அண்டியிருப்போர் (Dependents) என்ற வகையாகப் பிரிவுகள் அமைந்தன. வருமான (income) அடிப்படையில், (i) ஊதியம் பெறுவோர், (ii) போதிய ஊதியம் இல்லாததால் ஓரளவு அண்டியிருப்போர் (Partly dependents) (iii) முற்றிலும் அண்டியிருப்போர் (Complete dependents) என்ற வகையில் பிரிவுகள் கருதப்பட்டன ஆனால் 1961-ல், குறிப்பாக, 'பொருளாதார அடிப்படையில் உற்பத்தியோடு தொடர்புடைய' (economically productive) பல பணிகளில் ஈடுபட்டிருப்போர் பொதுவாகப் 'பணியாளர்' அல்லது 'உழைப்போர்' (workers) என்ற பிரிவில் இடம் பெற்றனர். அதற்கு, ஊதியம் பெற்றும் அல்லது பெறுமலும் இருக்கலாம். இத்தகைய பணிகளில் ஈடுபடாதோர், 'உழைப்போர் அல்லாத' பிரிவைச் சார்ந்தவராக கருதப்பட்டனர்.

(5) முதன்முதலாக, ஜம்மு-காஷ்மீர் மற்றும் சில பனிப் பிரதேசங்கள் ஆய்வில் இடம்பெற்றன. மேலும், இரு மொழிகள் கருதப்படாது, வேறு பல மொழிகள் பற்றிய (தாய்மொழி தவிர) அறிவும் கருதப்பட்டது.

(6) திருமணம் பற்றிய (marital status) விவரங்கள் பெறப்பட்டன. சமூகத்தில் சில பிரிவினர் (விலைமகளிர் போன்றோர்), இதன்மூலம் ஒரு குறிப்பிட்ட பிரிவில் இடம்பெற வழி ஏற்பட்டது.

(7) வேறிடத்தை விட்டுப் பெயர்ந்து வந்த தன்மை (migration, கிராமம் மற்றும் நகரம் என்னும் பிரிவுகளுடன்)

மற்றும் தங்கியிருந்த காலவரை பற்றிய விவரங்கள் கருதப்பட்டன.

சில விவரங்கள் (1961) :

- ☐ இந்திய மக்கட் தொகை : 4, 39, 235, 082.
- ☐ ஆண்கள் : 2, 26, 293, 620.
- ☐ ஏனையோர் பெண்கள் ஆண்; பெண் விகிதம் :—941 பெண்கள் (1000 ஆண்களை கருதி)
- ☐ ஒவ்வொரு ஆண்டும் மக்கட் தொகை பெருகும் வீதம் : 2·2%
- ☐ கிராமப்புறத்தில் வசிப்போர் சதவீதம் : 82·2%
- ☐ எழுதப்படிக்க தெரிந்தோர் : 24.03%

(5) 1971 மக்கட் கணிப்பு ஆய்வு
(1971 Population Census.)

நம் நாட்டின் சமூக மற்றும் பொருளாதார வளர்ச்சிக்கு ஏற்ப, விரிவாகவும் தெளிவாகவும் மக்கட் தொகைக்கணிப்பு விவரங்கள் பெறுவதன் இன்றியமையாத தன்மையினைக் கருத்திற்கொண்டு, ஆய்வு முறைகளும், செயல்படும் விதமும் பெரிதும் மாற்றப்பட்டு இருப்பதை 1971-ல் நடைபெற்ற மக்கட் கணிப்பினில் கண்டறியலாம்:

முன்னோடியான பணிகள், 1967-ம் ஆண்டின் துவக்கத்திலேயே மேற்கொள்ளப்பட்டு, கணிப்பில் பயன்படுத்தப்படவேண்டிய கருத்துக்களும் (Concepts), செயல்முறைகளும் (Procedure) நன்கு விவாதிக்கப்பட்டன. முன்னோடி ஆய்வுமூலம் நடைமுறைச் சிக்கல்கள் ஏற்படும் அனுபவங்கள் பற்றிய தெளிவான முடிவுகள் பெறப்பட்டன. இவை அனைத்தும் 1968-ன் துவக்கத்திலேயே முடிவுபெற்றன. மேலும் ஒருமுறை கண்காணிப்பாளர்களும் (Superintendents), விவரகளைப் பயன்படுத்துவோரும் கூடி விவாதித்தபின்னர், திட்டமிடலின் இறுதி வடிவம் பெறப்பட்டது (1969).

பின்வரும் வினாப்பட்டியல்கள் (Schedules) கணிப்பில் பயன்படுத்தப்பட்டன :

(i) வீடுகள் பற்றிய விவரப்பட்டியல் (Houselist)
(ii) தனிநபர் ஆய்வுச்சீட்டு (Individual Slip)
(iii) மக்கட்தொகை குறிப்பு (தனிநபர் சீட்டுகளைக் கருதி) (Population Record)

(i) ஆனது, (அ) எல்லா வீடுகளையும் கருதி அமைக்கப் பட்ட முக்கிய வீட்டு விவரப் பட்டியலாகவும் (Main houselists), (ஆ) வசிக்குமிடமாகவும் அல்லாமலும் அமைந்த 'கட்டிடங் களின்' விவரப்பட்டியலாகவும் (Subsidiary establishment schedule for residential cum non-residential buildings) என்ற வகையில் இரு பிரிவுகளாய் அமைந்தது.

வீடுகள்பற்றிய விவரப்பட்டியல் :—

இதன்மூலம், ஒவ்வொரு கிராமம் மற்றும் நகரில் அமைந்தி ருக்கும் குடும்பங்கள் மற்றும் 'கணக்கெடுப்பு இல்லங்கள்' பற்றிய விரிவான விவரப் பட்டியல் கருதப்படுகிறது. தோராய அளவில் மக்கட் தொகையும் கருதப்படும். 1970-ல் இத்தகைய பட்டியல்கள் நடைமுறைப்படுத்தப்பட்டு, கணக்கெடுப்புத் தொகுதிகள் நிர்ணயிக்கப்பட்டன. இதைத் தெளிவுபடுத்திக் கொள்வதன்மூலம், முழு அளவில் இடப்பரப்புகள் கருதப்படவும் (complete coverage), சில விவரங்கள் விட்டுப்போகாமலும், திரும்ப வராமல் இருக்கும்படியும் அமைய முயற்சி மேற் கொள்ளப்பட்டது. இது, 1961-ல் அமைக்கபட்ட பட்டி யலின் அடிப்படையிலேயே உருவாக்கப்பட்டது. காலியாக இருக்கும் 'இருப்பிடங்கள்' 'குறிப்புக்கள்' (Remarks) என்ற பத்தியின்மூலம் அறியப்பட்டன. குடும்பத்தலைவர் தாழ்த்தப் பட்ட வகுப்பினராக அல்லது பிற்பட்டோர் பிரிவாக இருப்பின், அதனைத் தெளிவாகக் குறிப்பிட வசதி செய்யப்பட்டது. பயிர்த் தொழிலில் ஈடுபடும் குடும்பங்கள்பற்றியும் தனியாகக் குறிப்பிடப் பட்டு, இப் பட்டியலானது விவசாயக் குடும்பங்கள்பற்றிய மாதிரி ஆய்வுகட்கு பயன்படத்தக்க விதத்தில் அமைக்கப்பட்டது.

தொழில் நிறுவனங்கள்போன்ற பல இதர வழிகளில் பயன் படுத்தப்படும் 'இல்லங்களின்' தன்மையினைத் தனியாக அறிய, 'நிறுவனம்பற்றிய விவரப் பட்டியல்' உருவாக்கப்பட்டது. அதன்மூலம், உற்பத்தி, வர்த்தகம் போன்ற பல நிறுவன அமைப்புகளில் விவரங்கள் பெற வழி ஏற்பட்டது, ஒவ்வொரு நிறுவனத்திலிருந்தும் (establishment) பின்வரும் பல்வேறு விவரங்கள் பெறப்பட்டன :

நிறுவனத்தின் பெயர் ; அரசுத்துறையினைச் சார்ந்ததா அல்லது தனியார் துறையா ; கூட்டு முயற்சியா (cooperative); சென்ற கால அளவில் சராசரியாக பணிப் புரிந்தோர் ; சிறு அளவில் அமையும் தொழிலா (household industry) ; பதிவு செய்யப்பட்ட தொழில் நிறுவனமா (Registered factory) ; பதிவு செய்யப்படாததா ; உற்பத்தி பொருட்களின் தன்மை ; வாங்கப்

பட்ட/விற்கப்பட்ட பொருள் அளவு ; வேறு ஏதாவது வழியில் பயன்படும் நிறுவன அமைப்பா? (அதாவது பள்ளிக்கூடம், மருத்துவமனை, சினிமா தியேட்டர், ஓட்டல் போன்றவைகட்கு) போன்றவை.

இத்தகைய பட்டியல்மூலம், பெரிதும் பயன்பெறும் விவரங்கள் பெறப்படுகின்றன. குறிப்பாக, ஒழுங்கானமுறையில் அமையாத அமைப்புகளைப்பற்றி அறிய இயலுகிறது. வெவ்வேறு இயல்புகளிடையே உள்ள தொடர்புகளும் அறியப்படுகின்றன. இதன்மூலம், மாதிரிமுறை ஆய்வுகள் மேற்கொள்ளப்பட முழு அளவில் 'சட்ட' (Frame) வரையறை கிடைக்கப்பெறுகின்றது. தொழில், வர்த்தகம் போன்ற பிரிவுகளில் உள்ள நிலையினை நன்கறிய இது உதவுகிறது.

தனிநபர் ஆய்வு (Individual slip)

இதன் வினாப்பட்டியல் தனியாக இணைக்கப்பட்டுள்ளது. இதில் மொத்தமாக 17 வினாக்கள் இடம் பெற்றன. அவற்றைச் சுருக்கமாகக் கூறலாம் :—

(1) நபரின் பெயர் தெளிவாகக் குறிப்பிடப்பட்டது. புதிதாகப் பிறந்த குழந்தையாக இருப்பின், 'குழந்தை' என்று மட்டும் குறிக்கப்பட்டது. மேலும், கருதப்பட்ட நபர் குடும்பத் தலைவருடன் எந்த விதத்தில் உறவுபெற்றிருக்கிறார் (relationship) என்பதும் விவரமாகக் குறிக்கப்பட்டது. தங்கும் விடுதிகள், உணவு விடுதிகள் போன்ற இருப்பிடங்கள் கருதப்படுகையில் அவற்றை நிர்வகிக்கும் அதிகாரிகள் 'தலைவர்' என்று கருதப்பட்டு (head), விவரங்கள் சேகரிக்கப்பட்டன. 'விருந்தாளி' 'பணியாளர்' (employee) என்பன போன்ற குறிப்புகள் அமைக்கப்பட்டன.

(2) மக்கள்தொகையானது, ஆண்–பெண் என இருவகையாகப் பிரிக்கப்பட்டது. M எனில் ஆண் எனவும், F எனில் பெண் எனவும் தனியாகக் குறிக்கப்பட்டு, பாடுபாடு செய்யப்பட்டது. அடுத்து, வயதினை (age) பொறுத்த அளவில் கடந்த பிறந்த நாளின்போது 'முழுமைபெற்ற ஆண்டுகள்' (Completed) குறித்துக்கொள்ளப்பட்டது. ஒரு வயதுக்கு கீழாக உள்ள நிலையில், 'சிறுகுழந்தை' (infant) என்று குறிப்பிடப்பட்டது.

(3) திருமண நிலையில், (marital status), N. M. (never-married), என்பது எப்போதுமே மணமாகாத நிலையையும் M

(married) என்பது ஆய்வுக்காலத்தில் திருமணம் ஆன நிலையையும் குறித்தன. W (widowed) விதவை நிலை; S திருமண முறிவு நிலை (divorced). இதில் மற்றொரு கேள்வியான 'தற்போது மணமாகியிருக்கும் பெண்களை'ப் பற்றிய அளவில், பெண்களிடமிருந்து விவரங்கள் பெற முயற்சிக்கப்பட்டது. திருமணத்தின் போது, வயது, கடந்த ஆண்டில் பெற்ற குழந்தை விவரம் போன்றவை மூலமாக, மக்கள்தொகையோடு தொடர்புடைய 'வீரியத் தன்மை' (fertility) பற்றிய விவரம் பெறப்பட்டது. 'மக்கட்தொகைப் பெருக்க' (population explosion) பிரச்சினை மிக முக்கியத்துவம் கொடுக்கப்படுகிறது.

(4) பிறந்த இடம் பற்றி, கிராமம் அல்லது நகர்ப்புறம் என்று இரு பிரிவுகள் அமைந்தன. ஆய்வுப் பகுதி (enumeration place) தவிரப் பிறிதோரிடத்தில் பிறந்ததாக இருப்பின், தெளிவாகக் குறிக்கப்பட்டது. விவரம் சரியாகத் தெரியவில்லை எனில், 'தெரியாது' (not known) என்று குறிப்பிடப்பட்டது. அதாவது, ஆய்வு நடைபெறும் பகுதி, அல்லாத பிற பகுதி என்று பிரிக்கப்பட்டு, ஒவ்வொரு பகுதியிலும், கிராமம்/நகரம், மாவட்டம், மாநிலம்/நாடு என்ற படியான விவரங்கள் தகுந்த குறியீடுகள் (codes) மூலம் குறித்துக் கொள்ளப்பட்டன.

அப்பகுதியில் இல்லாது, வேறு இடத்தில் வசிப்பவராக இருப்பின், தெளிவாகக் குறிப்பிடப்பட்டது. மேலும், ஆய்வு நடைபெறும் கிராமப்புறம் அல்லது நகரில் வசிக்கும் கால அளவு (duration of Residence), முழு ஆண்டுகளாகக் குறிக்கப்பட்டது. '0' என்பது ஓராண்டுக்குக் குறைந்தக் காலத்தை குறித்தது.

(5) மதங்கள் பற்றிய வரையில், இந்து (H), முஸ்லிம் (I), கிறித்தவர் (C), சீக்கியர் (S), ஜைனர் (J), புத்தமதம் (B), இதரவகை என்று பிரிவுகள் கருதப்பட்டன. அடுத்து, பிற்பட்ட வகுப்பு, தாழ்த்தப்பட்ட வகுப்புபற்றி விவரங்கள் அறியப்பட்டன.

(6) எழுதப்படிக்கத் தெரிந்த நிலை ஆராயப்பட்டது. 4 வயதுக்குக் கீழாகவுள்ள குழந்தைகள், எழுதப்படிக்கத் தெரியாத வகையில் இடம் பெற்றனர். மேலும், வெவ்வேறு கல்வி நிலைகள் கண்டறியப்பட்டன. ஒருவரது, பெற்ற அதிகமான கல்வித் தகுதி குறிக்கப்பட்டது. மொழியறிவு பற்றி, தாய்மொழிகள் குறிக்கப்பட்டு, இதர மொழிகள் பற்றிய அறிவும் தனியாகப் பெறப்பட்டது.

(7) முக்கிய பணிகளை (main activ'ty), (i) உழைப்பாளர் (ii) அல்லாதோர் (non-workers) என்று இரு பெரும் பிரிவில் கருதினர். இதன்படி, 'பொருளாதார உற்பத்திபற்றிய ஆக்கப் பணியில் இடம் பெறும்' உழைப்பாளர்கள் முதல் பிரிவில் இடம் பெற்றனர்.

C என்பது பயிர்த்தொழில் (cultivation) புரிவோரையும்; AL என்பது விவசாயப் பணியாளரையும் (agri-worker); H என்பது வீட்டு வேலைகள் (household activities) புரிவோரையும், ST என்பது மாணவரையும், HHI எனில் வீடுகளில் நடைபெறும் சிறு தொழில்களையும் OW இதரப் பணிகள்; R ஓய்வு பெற்ருர்; B பிச்சைக்காரர் போன்ருேர்; O–மற்றவை என்று பலவாருகவும் விரிவான பாகுபாடு அமையப்பெற்றது.

வேலை செய்யும் இடம் (கிராமம்/நகரம்), நிறுவனம், தொழிலகம், வர்த்தக அமைப்பு, விற்பனை அகம், அலுவலகம், (மத்திய அல்லது மாநில) அரசினைச்சார்ந்த நிலை, பணியின் தன்மை, தொழில் நிலை, போன்ற பல்வேறு விவரங்களும் தக்க பிரிவுகளில் அமைக்கப்பட்ட வினுக்களின்மூலம், உறுதி செய்யப்பட்டன.

வெவ்வேறு பணிகளில் ஈடுபடுவோர் தெளிவாகத்தமது பணி நிலையினையும் தன்மையையும் குறிப்பிடுமாறு கவனித்துக் கொள்ளப்பட்டது.

எனவே 1971–ல் பயன்படுத்தப்பட்ட தனி நபர் ஆய்வுச் சீட்டானது, மக்களைப்பற்றி பொருளாதார, சமூகப், பிறப்பு இறப்பு பற்றிய, பல்வேறு வகையான விவரங்களையும் பெறத் தக்க விதத்தில் விரிவாகவும் தெளிவாகவும் உருவாக்கப்பட்டது.

வெவ்வேறு ஆய்வுகளுக்கும் பயன்படத்தக்க விதத்தில் மக்கள்தொகை குறிப்புக்கள் (population record) விவரமாக அமைக்கப்பட்டன.

மேலும் இந்தக் கணிப்பாய்வில், துளை அட்டை (punch card) மற்றும் மின்னியற்பகுப்புக் கணக்கீட்டுப் பொறிகள் (electronic computers) பெரும் அளவில் பயன்படுத்தப்பட்டு, பாகுபாடு மற்றும் இறுதி அட்டவணைப்படுத்தல் (sorting and tabulation) ஆகிய பணிகள் செம்மையாக நடத்தப்பட்டன. ஆய்வின் குறிப்பீட்டுக் காலம் (reference period) 1971, 'ஏப்ரல் ஒன்ரும் நாள் அதிகாலை' என்று நிர்ணயிக்கப்பட்டது. இந்த நாளில்

கருதப்பட்ட இருப்பிடங்களில் 'வசிக்கும்' நபர்கள் கணக்கெடுக்கப்பட்டனர். ஏப்ரல் 1 முதல் 3-ம் தேதி வரையில், மறு ஆய்வு விரிவாக மேற்கொள்ளப்பட்டது. இதன்மூலம், மிகவும் சரியாக (accurate) இருக்கும்படியாக விவரங்கள் பெறப்பட்டன. 1971, ஆகஸ்ட் 16-ம் நாள், பொதுப் பதிவு அதிகாரி (Registrar General of Census) மூலம், மக்கட்தொகை அளவு வெளியிடப்பட்டது (546,955,945). 1961 முதல் 1971 வரையிலான பத்தாண்டுக் காலத்தில், மக்கட்தொகை 10·8 கோடி அதிகரித்திருப்பதாகக் கணக்கிடப்பட்டது. (27·6%) இறப்பு வீதம் (mortality rate) அதிகரிக்காத நிலையில், குடும்ப நலத் திட்டங்கள் பெரிதும் முக்கியத்துவம் அளிக்கப்பட்டு, தகுந்த நடவடிக்கை எடுக்கப்பட்டு, மக்கட்தொகைப் பெருக்கத்தைக் கட்டுப்படுத்த முயற்சிகள் மேற் கொள்ளப்பட்டு வருகின்றன.

நகர்ப்புற மக்கட்தொகை அதிகரித்தும்கூட, பெரும்பான்மையான மக்கள் (80%) கிராமப் புறங்களில் வசிப்பதாகவே அறியப்படுகிறது. தமிழ்நாட்டில் நகர்ப்புற மக்கள் எண்ணிக்கை 30·28% என்று அறியப்பட்டது. இது, இரண்டாவது இடம் பெறுகிறது (முதலாவது: மகாராஷ்டிரம்: 31·20%). உலகிலேயே பெரிய அளவினவாகக் கருதப்படும் 25 பெருநகர்(cities)களுள், கல்கத்தா, பம்பாய், டெல்லி போன்று மூன்று நகர்கள் இந்தியாவில் இடம்பெறுவது குறிப்பிடதக்கது. பின்வரும் சில புள்ளி விவரங்கள் 1971 மக்கட் கணிப்பின் அடிப்படையில் பெறப்பட்டவை ஆகும்:

1971 இந்திய மக்கட் கணிப்பு ஆய்வில் பயன்படுத்தப்பட்ட தனிநபர் ஆய்வுச் சீட்டு (Individual Slip)

அட்டை எண்.........
சீட்டு எண்.........

[இரகசியமானது]

இடக் குறியீடு (Location Code) குடும்ப எண்:

1. பெயர்......
2. குடும்பத்தலைவருடன் உறவு......
3. இனம்...... (Sex)
4. வயது ......
5. திருமணம் ஆனவரா ...
6. தற்போது மணமான பெண்களுக்கு மட்டும்:
   (அ) திருமணத்தின் போது வயது ......
   (ஆ) கடந்த ஓராண்டில் ஏதேனும் குழந்தை பிறந்ததா ? W...

7. பிறப்பிட விவரம் (Birth place):-
(அ) பிறந்த ஊர்......
(ஆ) கிராமம்/நகரம்....
(இ) மாவட்டம்......
(ஈ) மாநிலம்/நாடு......

8. இறுதியாக வசித்த இடம் :
(அ) வசித்தஊர்......
(ஆ) கிராமம்/நகரம்......
(இ) மாவட்டம்......
(ஈ) மாநிலம்/நாடு.....

9. ஆய்வு நடைபெறும் கிராமம் அல்லது நகரில் வசித்த கால அளவு :
....
....

10. மதம்......
(Religion)

11. தாழ்த்தப்பட்ட வகுப்பு: ...... அல்லது பிற்பட்ட வகுப்பு

12. எழுதப்படிக்கத் தெரிந்தவரா!
(L அல்லது O) :

13. கல்வி நிலை......

14. தாய் மொழி......

15. அறிந்த வேறு மொழிகள்......

16. முக்கிய அலுவல் விவரம்:
(Main Activity)
(அ) விரிவான பிரிவு
(i) வேலையாள் (C,AL,HHI, OW)
....
(ii) அல்லாதவர் (H,ST,R,D,BI,O)
....
(ஆ) பணி புரியும் இடம்..... (கிராமம் அல்லது நகரின் பெயர்)
(இ) நிறுவனத்தின் பெயர்... (Establishment)
(ஈ) தொழில் அல்லது வர்த்தகம் அல்லது பணி அல்லது ஊழியம் இவற்றின் தன்மை (Nature)
....
....
(உ) பணியின் விவரம் (Description of work)......
....
....
(ஊ) வேலையாளர் எனில், எப் பிரிவு (class):
....

17. துணை அலுவல் விவரம் (Secondary Activity) :
(அ) விரிவான பிரிவு :......
(C, AL, HHI, OW)
(ஆ) பணி புரியும் இடம்:...... (கிராமம் அல்லது நகரின் பெயர்)
(இ) நிறுவனத்தின் பெயர்......
(ஈ) தொழில், வர்த்தகம், பணி, அல்லது ஊழியம், இவற்றின் தன்மை......
....
....
(உ) பணி விவரம்......
(Description of work)
....
....
(ஊ) வேலையாளர் எனில், பிரிவு...... (class)

1971 மக்கட் கணிப்பில் சில விவரங்கள்

- மக்கட்தொகை : 54,69,55,945
- ஆண்கள் : 28,30,55,987
  (ஏனையோர் பெண்கள்)
- ஆண் பெண் விகிதம் : 932 பெண்கள் (1000 ஆண்களைக் கருதி)
- பத்து ஆண்டுகளில்
- மக்கட்தொகை அதிகரிப்பு வீதம் : 24·66%
- எழுதப்படிக்கத் தெரிந்தோர் சதவீதம் : 29·34
- கிராமங்களில் வசிப்போர் சதவீதம் : 80·13%
- உழைப்போர் வீதம் (Workers) : 33·54%
- பயிர் செய்வோர் வீதம் : 42·87%
- விவசாயத் துறையில் உழைப்போர் வீதம் : 25·76%
- இதரப் பணியாளர் : 31·37%

இதன்மூலம் நாட்டு மக்களின் பல்வேறு இயல்புகள்பற்றிய ஒரு மதிப்பீடு பெறப்படும். இந்த முடிவுகள் பல்வேறு பிரச்சினைகளைக் குறித்து நடத்தப்பெறும் ஆய்வுகளிலும் பெரும் அளவில் உதவி புரிகின்றன.

முப்பது கோடி சதுர கிலோ மீட்டர் (km) பரப்பில் சுமார் 60 கோடி மக்களைக் கருதி, பெரும்அளவில் திட்டமிட்டு விரிவான கணிப்பாய்வு மேற்கொள்ளுவது என்பது, நடைமுறையில் பலவகையான இடர்ப்பாடுகளுடன், சிக்கல்களுடனும் அமைவது குறிப்பிடத்தக்கது. மக்களுடைய ஒத்துழைப்பு இல்லாமல் இதனை வெற்றிகரமாகத் திட்டமிட்டு நடத்தி முடிப்பது நினைக்க இயலாத செயலாகும். நடைமுறையில் சிலபிழைகள், தவறுதல்கள் ஏற்படுவது தவிர்க்க இயலாததொன்றுகும். சில இடர்ப்பாடுகளை இங்கு தொகுத்து அளிப்போம்:

(6) மக்கட் கணிப்பில் சில நடைமுறை இடர்ப்பாடுகள்

(1) மக்களின் அலட்சியப்போக்கு: (Indifference)

அனைவரது பூரண ஒத்துழைப்பின் விளைவே மக்கட்தொகை கணிப்பின் இறுதி வடிவமாக உருப்பெறுகிறது. இதன் நடை முறைகளைப் பற்றிச் சிறிதும் கவலைப்படாமல் மெத்தனமாக நடந்துகொள்ளும் மக்களின்மூலம் முடிவுகள் பெரிதும் பாதிக்கப்படுகின்றன. சரியான, துல்லியமான (Accurate) முடிவுகளைப் பெறுவதானது முழுவதும் பொதுமக்கள் தருகின்ற விடைகளை (Replies) பொறுத்து அமைகின்றது. இதனை நீக்கு வதற்கு, தகுந்தஅளவு விவரங்களைப் பெறுவது நாட்டின் பொரு ளாதாரத் திட்டங்கள் அமைத்திட எவ்வளவு தூரம் உதவி புரிகின்றது என்பதைப்பற்றி பொதுமக்களுக்கு நன்கு எடுத்து விளக்கி அறிவுறுத்தி அவர்தம் ஒத்துழைப்பினைப் பெறுவது அவசியமான ஒன்றாகும். பத்திரிகைகள், படக்காட்சிகள், வானொலிகள்மூலம் இப்பணிகள் மேற்கொள்ளப்பட வேண்டும்.

(2) போதிய பயிற்சியற்ற ஆய்வாளர்மூலம் இன்னல்கள்

பல்வேறு ஆய்வுப் பணியாளரும் தற்காலிகமாக நியமிக்கப் படுகின்றனர். அவரது திறமையினை (efficiency) பொறுத்தே பல்வேறு பிறப்பு இறப்புபற்றிய விவரங்கள் (Demographic) மற்றும் இதனோடு தொடர்புடைய இதர விவரங்கள் ஆகிய வற்றின் நன்பகத்தன்மையும் ஏற்புத்தன்மையும் அமைந் திருக்கும் என்று கூறலாம். இயல்பான அலுவலக நேரத்துக்கு அப்பால் அமையும் (Out-of office hours) காலங்களில் கணக்கெடுப்பு நடத்தவேண்டியிருக்கிறது. எனவே முழுக்கவனம் செலுத்தி விவரங்களைப் பெற இயலாமல் போகலாம். அவர்தம் பொறுப்பும் அதிகமானதாக ஆகிறது. சிறிய கால அளவினதான பயிற்சிதான் அளிக்கப்படுகிறது. எனவே சில கருத்துக்களை ஆய்வாளர்கள் தெளிவாக அறியாமலும், வெளிப் படுத்த இயலாமலும், புரிந்துகொள்ளாமலும், சில வினாக்களின் உட்கருத்தைச் சரிவர உணராமலும் இருக்கின்ற நிலைகள் ஏற்படுகின்றன. சில நடைமுறைகளை பற்றிய விவரங்களை, கொடுக்கப்படும் விவரக்குறிப்பு ஏடுகளினின்று பெரும்பாலான ஆய்வாளகர்ள் நன்கு படித்து அறிந்துகொள்வதில்லை. எனவே நம்பத்தக்க விவரங்களை அறிய இயலாமற் போகிறது. தரப்படும் ஊதியமும் போதிய அளவில் இருப்பதில்லை. எனவே தகுந்த நடவடிக்கைகள்மூலம் இக்குறைப்பாட்டினை நீக்கலாம்.

(3) ஒவ்வொரு கணிப்பாய்விலும், கருதப்படும் இடப்பரப்பு (பூகோள அடிப்படையில்) மாறுபடும் தன்மை உடையதாகக்

காணப்படுகிறது இதைத்தவிர்க்க வழிமுறைகள் மேற்கொள்ளப் படின், ஒழுங்கான முறையில் விவரங்கள் பெறலாம். மேலும், சில பிரிவினரைப் பற்றிய முழுவிவரங்கள் (உதாரணமாகப் பிச்சைக்காரர்கள், அங்குமிங்கும் திரிவோர் போன்ற பலர்) அறியப்படாமல் போகலாம். இதற்கும் தக்கவழிகள் மேற் கொள்ளப்படுதல் நல்லது.

(4) சில வினுக்களைப் பொறுத்த அளவில், தெளிவான கருத்துக் களும் வரையறைகளும் தரப்படுவதில்லை. தெளிவற்ற நிலையில் வினுக்கள் அமைந்திடின், பதில் தருவதில் சிரமம் ஏற்படுகின்றது, உதாரணமாக, 'எழுதப் படிக்கத் தெரிந்த' நிலைபற்றி அறிய வேண்டின், தெளிவாக அந்நிலையினை விளக்கவேண்டும். அதாவது, 'ஏதேனும் மொழியில், ஒரு கடிதமாவது எழுதத் தெரிந்திருக்கும் வகையில் அறிவு பெற்றிருக்கவேண்டும்' என்று வரையறுக்கப்படின், பதில் பெறுவது எளிதாகிறது. இதேபோல் மற்ற நிலைகளும் கருதப்படவேண்டும்.

(5) முக்கியமானதாக அமைவது, 'வயதுபற்றிய விவரங்களைப் பெறுவதாகும். இக்கருத்து, ஒவ்வொரு கணிப்பாய்விலும் மாற்றப்பட்டு வந்திருக்கிறது. (இறப்பு வீதம்பற்றிய மதிப் பீட்டிற்கு இவ்விவரங்கள் பெரிதும் பயன்படுகின்றன. (உதாரணமாக, ஆயுள் பாதுகாப்புத் திட்டம் (Life Insurance Plans) பற்றிய மதிப்பீடுகள். வயது பற்றிய விவரங்கள் சரிவரக் கொடுக்கப்படுவதில்லை. மக்களின் அறியாமை, மெத்தனப் போக்கு, எழுத்தறிவின்மை போன்றவற்ருல் சரியான வயது விவரங்களைப் பெற இயலாமல் போகிறது. பெண்கள் தம் வயதினைக் குறைவாகக் கொடுக்கவும், அல்லது கொடுக்க மறுக் கவும் நேரிடுகிறது. தவருன விவரங்கள் குறிக்கப்படுகின்றன. ஆய்வாளின் தீர்ப்புப்படி விவரங்கள் அமைகின்றன. கிராமப் புரங்களில் இச்சிக்கல் பெரிய அளவில் ஏற்படுகிறது. சில மூட நம்பிக்கைகள் காரணமாகவும் விவரங்கள் மறுக்கப்படுகின்றன. தகுந்த உத்திகளைப் பயன்படுத்தி, இவ்விவரங்களைப் பெறுவது ஆய்வாளரின் சாமர்த்தியத்தைப் பொறுத்தது. இவ்விவரங்கள் மூலம், வெவ்வேறு வயதுப்பிரிவுகளில் இடம்பெறும் ஆண் பெண் விவரம், கல்வியறிவுநிலை, மணமான விவரம், பொருளா தாரநிலை போன்ற பலவற்றையும் விரிவாகக் கருதவேண்டியிருப் பதால் இதில் தனிக் கவனம் செலுத்தவேண்டியமையாதாகிறது.

மக்கட் கணிப்பின் தேவையும் பயனும் (Importance & Utility)

நாட்டில் வாழும் பல்வேறு மக்களைப் பற்றிய விவரங்கள் விரிவான அளவில் பெறப்படுவதன்மூலம் சமூக, பொருளாதார, அரசியல் இயல்புகள் பெரும் அளவில் அலசி ஆராயப் படுகின்றன. சிலவற்றை மட்டும் தொகுத்து இங்கு அளிப்போம்:

(1) நாட்டின் வர்த்தகம், வணிகம் மற்றும் தொழில்துறை போன்ற பல்வேறு துறைகளிலும் மக்கட்தொகைக் கணிப்பு முடிவுகள் பெரிதும் உதவி புரிகின்றன. சில பொருட்களை (பணிகளும் கூட) உற்பத்தி செய்கையில், அதற்குக் கிடைக்கப் பெறும் உழைப்பாளர்கள் எண்ணிக்கை, இனம், வயதுபோன்ற வற்றிற்கு ஏற்ப அவர்தம் பாகுபாடு (classification) போன்ற வற்றை முன்கூட்டியே அறிந்துகொள்வது, செயல்முறைத் திட்டத்திற்குப் பெரிதும் உதவும். மேலும், நாடுமுழுவதும் பரவி யுள்ள நுகர்வோர்கள் (consumers) தன்மை, அவர்தம் பொருளா தாரச் சூழ்நிலை, ஈடுபடும் பணிகள், கல்விநிலை போன்ற பல விவரங்களை அறிந்து, வெவ்வேறு பொருட்களின் தேவை அளவுகள் மதிப்பிடப்படுவது, விற்பனையாளரின் செயல் முறைக்கு மிக அவசியமாகிறது. வங்கிகள், வர்த்தக நிறுவனங்கள், ஆய்வுக் காப்பகங்கள் (Insurance companies) போன்றவை மக்களைப்பற்றிய எல்லா விவரங்களையும் பெற்று, மதிப்பீடுகளை அடைய இக் கணிப்பானது மிகமிக தேவைப்படுவதொன்று. எதிர்பார்க்கும் ஆயுளினை (expectation of life) மதிப்பிடுவதன் மூலம், சரியான திட்டங்களை மேற்கொள்ள, ஆயுள் காப்பகத்தி னால் இயலுகிறது.

(2) தற்போது அதிகம் முக்கியத்துவம் அளிக்கப்பட்டு வரும் 'மக்கள் இயல்' (demography) பற்றிய ஆராய்ச்சிகளுக்கு மக்கட் கணிப்பாய்வு விவரங்கள் மிகவும் உதவுகின்றன. நகரச் சுற்றுப்புறச் சூழ்நிலைகளை அறிவதன்மூலம், சமூக நிலைகளைக் கண்டறிய முடிகிறது. சில சட்ட வரையறைகளை ஏற்படுவதற்கு, இத்தகைய ஆய்வுகள் அடிப்படையாக அமைகின்றன. சில சட்டங்களை அல்லது பிரேரணைகளை நிறைவேற்றுவதில், புள்ளி விவரங்கள் பெரிதும் தேவைப்படுகின்றன, அதாவது, மக்கள் அவை (legislature) போன்ற அமைப்புகள் சமூக நலனைக் கருத்திற்கொண்டு செயல்பட, சமூக நிலைகளைப் படம் பிடித்துக் காட்டும் விவரங்கள் அவசியமாகின்றன. தொகுதிகளில இருக்கும் மக்கட்தொகைக்கு ஏற்ப, பிரதிநிதித்துவம் அளிக்கப்படும். அதேபோல, சமூக நல முன்னேற்றத் திட்டங்கள் அனைத்தும் மக்களைப் பற்றிய விவரங்களைக்கொண்டே மேற்கொள்ளப்பட வேண்டியதாகிறது.

(3) மக்களின் பொருளாதார நிலை பற்றிய அறிவு மிக இன்றியமையாததாகும், செய்யும் தொழில்களுக்கு (occupations) ஏற்ப அமைந்திடும் பாகுபாடானது (classification) அதிகக் கவனத்துடன் அமைக்கப்படுகிறது. பொருளாதார முன்னேற்றத்திற்கு மேற்கொள்ளப்படும் நாட்டின் பல்வேறு திட்டப் பணிகள், பெறப்படும் முழு விவரங்களின் அடிப்படையில் அமைதல் வேண்டும். வேலையில் இருப்போர் நிலையினையும் (employment), வேலையின்மை நிலையினையும் (unemployment) பற்றி அறிந்துகொள்வதன்மூலம், தக்க நடவடிக்கைகள் எடுக்க அரசினரால் இயலுகிறது. இதோடன்றி, மக்களின் கல்வியறிவு நிலை பற்றிய மதிப்பீடுகளின்மூலம், கல்வித் துறையில் தக்கபடி மாற்றங்கள் ஏற்படுத்தவும், நாட்டின் தேவைகட்கு ஏற்ப, அடிப்படைக்கல்வி, உயர்கல்வி, தொழிற்கல்வி போன்றவற்றைத் திட்டமிட்டு அமைக்கவும் அரசினருக்கு வழியேற்படுகிறது.

எனவே, மக்கட் கணிப்பின் முடிவுகள் விரிவான அளவில் பயன்படுத்தப்பட்டு, பரந்த நோக்கம் கொண்டனவாக அமைவதைக் காணலாம். இதனால்தான், பல நாடுகளில் தற்சமயம் மக்கட்கணிப்பு ஆய்வுகள் நடத்தப்படவும், மக்களியல் பற்றிய வெவ்வேறு பிரச்சினைகளைத் தீர்க்கவும் பெரும் அளவில் முயற்சிகள் மேற்கொள்ளப்பட்டு வருகின்றன. இந்தியப் புள்ளி விவரங்களுள் (Indian Statistics), மக்கட்தொகைபற்றி அமையும் விவரங்களும் முடிவுகளும் ஒரு முக்கிய இடத்தைப் பெறுகின்றன என்று கூறிடின் மிகையாகாது.

மக்கள் இயலில் (Demography) மிக முக்கியமான, 'பிறப்பு—இறப்பு' புள்ளிவிவரங்கள் திரட்டப்பட்டு ஆராயப்படும் சில வழிமுறைகள் அடுத்த பகுதியில் இடம்பெறும்.

மாதிரிப் பதிவீட்டுத் திட்டம்
(Sample Registration Scheme)

பிறப்பு—இறப்பு புள்ளிவிவரங்கள்: (VITAL STATISTICS)

*** இந்திய நாட்டின் மாதிரிப் பதிவீட்டுத் திட்டம் :**
(Sample Registration Scheme in India)

ஒரு நாட்டின் நல்வாழ்வு, பொருளாதாரம், சமூக முன்னேற்றம் ஆகிய பல்வேறு துறைகளில் விரிவான அளவில்

* Ref : Vital Statistics Division, Office of Registrar General, Ministry of Home Affairs, India (1972)

திட்டங்களைத்(Planning) தீட்டுதற்கு, அதனோடு மிகவும் தொடர்பு கொண்ட மக்கட்தொகை அளவு (size), அதன் பரவல் தன்மை (distribution), பல்வேறு வகை இயல்புகளின் கூட்டுத் தன்மை (composite nature) போன்ற பலவகை விவரங்களை அறிந்திடுவது இன்றியமையதாக ஆகிறது. மக்கட்தொகை பற்றிய இவ் விவரங்கள் அனைத்தும், வீரியத்தன்மை (fertility), இறப்பு வீதம் (mortality), இடம்விட்டு இடம்மாறும் மக்கள்வீதம் போன்ற பல நிகழ்ச்சிகளைப் பொறுத்து அமைவதை உணரலாம். உதாரணமாக, பங்களாதேஷ் பிரச்சினை காரணமாக நம்நாட்டில் அடைக்கலம் பெற்றோர் (refugees) எண்ணிக்கைமூலம், ஏற்கெனவே இருந்த மக்கட்தொகை மிகவும் அதிகரித்து இருக்கக் கண்டோம். அதற்குத் தக்க அளவில் பொருளாதார, சமூகத் திட்டங்கள் அனைததுமே, பிறப்பு—இறப்பு விவரங்கள் பற்றிய புள்ளியியலின் (vital statistics) தன்மையினைப் பொறுத்து மாறுபடுவது குறிப்பிடத்தக்கதாகும். இது பொதுவான புள்ளியியலில் ஒரு முக்கியமான துறையாக (field) அமைகிறது. இதில் ஆராயப்படுவது, மக்கட்தொகை தேற்றிய பல்வேறு பிரச்சினைகளாகும்.

நாட்டின் பிறப்பு இறப்பு பற்றிய புள்ளிவிவரங்களை (vital statistics) மூன்று விரிவான வழிமுறைகள்மூலம் பெறக் காணலாம். அவையாவன :

(i) பொதுவான பதிவு முறை (Civil Registration)

(ii) மாதிரி ஆய்வு முறை (Sample Survey)

(iii) பகுப்பாய்வு முறை (Analytic Method)

ஒரு இடத்தில் உள்ள அமைப்பின் மூலம் (agency), நிகழும் பிறப்புக்கள் மற்றும் இறப்புக்கள்பற்றிய முழு விவரங்களைத் தவறுது தொடர்ந்து பதிவு செய்வது முதல் வழி முறை. உதாரணமாக, பொதுமக்கள் அனைவரும் தத்தம் இல்லங்களில் நிகழும் பிறப்பு—இறப்பு பற்றிய விவரங்களை அதற்கேற்ற அதிகாரியிடம் பதிவு செய்து வைக்கும்படி கோரப்படுகின்றனர். அது அவர்தம் கடமை எனவும் வகுக்கப்படுகிறது. இம்முறை மூலம், நாட்டின் இராணுவத் துறையினைச் சாராத ஏனைய மக்களது விவரங்கள் அனைத்துமே பெறப்படுகின்றன.

முதல் வழிமுறையானது தக்கபடி திறமையாகச் செயல்படவில்லை எனில், பிறப்பு இறப்புபற்றிய விவரங்களைத் திரட்டும் பொருட்டு மாதிரி ஆய்வு முறைளானது மேற்கொள்ளப்பட்டு,

தேவைப்படும் விவரங்கள் பெறப்படுகின்றன. இந்த நோக்கத்திற்கேற்ற பிரத்தியேக ஆய்வுகள் நடத்தப்படுகின்றன.

முதல் இரு வழிமுறைகளும் பயன் தராது அமைந்திடும் நிலையில், பிறப்புவீதம் மற்றும் இறப்பு வீதம் ஆகியவற்றினைக் கணிக்க, தொடர்ந்து எடுக்கப்படும் இரு மக்கட் கணிப்புகள் பயன்படுத்தப்பட்டு, முடிவுகள் பெறப்படுகின்றன. இது பகுப்பு அல்லது பகுப்பாய்வு முறை எனக்கருதப்படும். உதாரணமாக, 1961 மற்றும் 1971 என்ற இரு ஆண்டுகளிலும் நடத்தப்பட்ட மக்கட் கணிப்பு ஆய்வுகளின் (Population Census) முடிவுகளைப் பயன்படுத்தி, இந்திய நாட்டின் பிறப்பு மற்றும் இறப்பு வீதங்கள் கணக்கிடப்படுகின்றன.

1971-ல் பதிவு முறை (Registration Scheme) மூலமாக, பின்வரும் விவரங்கள் மதிப்பிடப்பட்டன:

	கிராமப்புறம் (Rural)	நகர்ப்புறம் (Urban)
பிறப்பு வீதம் (திருத்தப்படாதது)	39	30
இறப்பு வீதம் (திருத்தப்படாதது)	17	10

(1000-க்குப் பெறப்பட்டவை)

கிராமங்களையும் நகரங்களையும் ஒருங்கு இணைத்துப்பெற்ற மதிப்புகள் 37,16 என்று முறையே அமைந்தன.

மாதிரி ஆய்வு :

தற்போது, இந்த ஆய்வு முறையானது இரு குறிப்பிட்ட வழிகளில் செயல்படுகிறது. முதலில், ஏற்றுக்கொள்ளதக்க கால இடைவெளிகளில் திருப்பப்படும் (repeated) ஆய்வுகளை நடத்தி முடிவுகள் பெறப்படும். இரண்டாவதாக, மாதிரியாகத் தெரிவு பெற்ற ஒரு இடப் பரப்பில் (area) தொடர்ந்து பதிவுகள் நடத்தப்படுகின்றன. இவற்றிற்கு வெவ்வேறு காலங்களில் நடத்தப்படும் ஆய்வுகள் பெரிதும் உதவி புரிகின்றன. இந்தியாவில், இவ்விரு முறைகளையும் இணைத்த செயல்முறை கையாளப்பட்டு, மாதிரிப் பதிவீட்டு முறை (Sample Registration Scheme—S. R. S.) என்று கூறப்படுகிறது.

S. R. S திட்டமுறையானது, சமவாய்ப்பு அடிப்படையில் மாதிரிகளாக தேர்வுபெற்ற கிராமங்களிலும், நகர்த் தொகுப்புக் களிலும் (urban blocks) நிறைவேற்றப்படுகிறது. தெரிந் தெடுக்கப்பட்ட பரப்புப் பகுதிகளில், பிறப்பு—இறப்பு விவரங்கள் அனைத்தும் தொடர்ந்து பதிவு செய்யப்படுகின்றன. ஆண்டுக்கு இருமுறை நடத்தப்படும் ஆய்வுகள் (half-yearly surveys), இவ்விவரங்களுக்குப் பெரிதும் துணைபுரிகின்றன.

சுருங்கக் கூறின், ஒரு குறுகிய காலவரையில் நாட்டின் பிறப்பு—இறப்பு வீதங்களை (birth and death rates) பற்றிய நம்பத்தக்க தெளிவான மதிப்பீடுகளைத், தனித்தனியான அளவில் கிராமப்புறத்திற்கும், நகர்ப்புறத்திற்கும், பெறுவதே S. R. S-ன் அடிப்படை நோக்கம் என்று கருதலாம். (பிறப்பு வீதம் மற்றும் இறப்பு வீதம் இவை இரண்டினுக்கு இடையில் உள்ள வேறுபாடு மக்கட்தொகை அதிகரிப்பு வீதத்தின் மதிப்பீடாகப் பெறப் படுகிறது என்பது குறிப்பிடத்தக்கது.) பிறப்பு—இறப்பு விவரங்கள் எந்த நாட்டிலும் அடிப்படையான தேவையாக விளங்குகிறது. நாட்டின் மக்கட்தொகை கட்டுப்படுத்தப்பட எடுக்கப்படும் பல்வேறு முயற்சிகளும் இவற்றினைப் பொறுத்தே அமைகின்றன. திட்டத்தில் முக்கியப் பிரிவாக, 'குடும்ப நலத் திட்டம்' இடம் பெறுவது காணத்தக்கது. எனவே, இத்துறை யானது அடிப்படை முக்கியத்துவம் பெறுவது எளிதாக விளங்கிடும். நாட்டின் பொருளாதார முன்னேற்றத்தைப் பாதிக்கும் முக்கியமான ஒன்றுக,மக்கட்தொகைப் பெருக்கமானது அமைகிறது.

மாதிரித் திட்ட அமைப்பு (Sample Design)

(i) கிராமியப் பகுதிகளில் (Rural)

கிராமங்களில் மாதிரி ஆய்வு நடத்துதற் பொருட்டு, ஒரே ஒரு நிலையில் (uni-stage) அமைந்த எளிய சமவாய்ப்பின் அடிப் படையில் கருதப்பட்ட படுகை மாதிரி முறையானது (stratified simple random sampling) பின்பற்றப்பட்டது. ஒரு மாநிலத்தில் உள்ள மொத்த கிராமப்புறப் பரப்பும் (village area), முழுமைத் தொகுதியாகக் கொள்ளப்பட்டு, முதல்நிலை மாதிரிக் கூருக (primary sampling unit), ஒரு கிராமம் அல்லது கிராமத்தின் ஒரு பகுதி (2000-க்கு மேல் மக்கட் தொகை இருக்கும் நிலைகளில்) யானது கருதப்பட்டது. மாதிரிச் சட்ட வரையினை (frame), 1961-ல் மாவட்ட வாரியாகத் தொகுக்கப்பட்ட, கிராமங்கள்பற்றிய விவரங்கள் கொண்ட முழுப் பட்டியலானது (directory)

அளித்தது. மக்கட்தொகை அடிப்படையில் கிராமங்கள் பிரிவு செய்யப்பட்டு, பட்டியல் அமைக்கப்பட்டது. பின்வருமாறு ஒரு அமைப்பு உருவாக்கப்பட்டது:

முதல் பிரிவு :	மக்கட் தொகை 500-க்குக் கீழே.		
இரண்டாம் பிரிவு :	,,	500 முதல்	999 வரை.
மூன்ரும் பிரிவு:	,,	1000 முதல்	1999 வரை.
நான்காம் பிரிவு :	,,	2000-க்கு மேல் (2000 உள்பட).	

இயற்கை அமைப்பின் அடிப்படையில் (natural division), முதலில் ஒவ்வொரு மாநிலமும் படுகைகளாகப் (strata) பிரிக்கப்பட்டு மேற்கண்டபடி மக்கட்தொகை அடிப்படையிலும் மேலும் பிரிவு படுத்தப்பட்டன. 2000-க்குக்கீழாக மக்கட்தொகை உள்ள கிராமங்களிலிருந்து சமவாய்ப்பின் அடிப்படையில் தேவைப்பட்ட கிராமங்கள் தெரிவுசெய்யப்பட்டன. 2000-க்குமேல் மக்கட் தொகை அமையப்பெற்ற நிலைகளில், ஒவ்வொரு கிராமமும் ஏறக்குறைய சமஅளவினதாக (2000க்குமேல் போகாது) இருக்கும் விதத்தில் பல பகுதிகளாகப் பிரிக்கப்பட்டது. இத்தகைய பகுதிகள் (segments) கொண்ட சட்ட அமைப்பு (frame) உருவாக்கப்பட்டு, தேவையான அளவு பகுதிகள் சமவாய்ப்பு அடிப்படையில் தெரிவு செய்யப்பட்டன. மாதிரிக் கூறுகள் அனைத்தும், வெவ்வேறு படுகைகளில், படுகை மக்கள் எண்ணிக்கையுடன் வீதத் தொடர்பு கொண்ட வகையில், பங்கிடப்பட்டன. இம்முறையானது கிராமப்புறங்களைப் பொறுத்த அளவில்மட்டுமே கையாளப்பட்டது. இனி நகர்ப் புறங்களை எடுத்துக் கொள்வோம்:

(ii) நகரப்புறங்களில் (Urban)

பொதுவாக, நகர்ப்புற மாதிரி வடிவ அமைப்பானது இரு நிலைகளை (two stages) உடையதாக அமைகிறது. இருநிலைகளிலும் படுகைப் படுத்தும் முறை (stratification) கடைப் பிடிக்கப்படுகிறது. நகரங்கள் (towns) முதல்நிலைக் கூறுகளாகவும், கணிப்பு ஆய்வுத் தொகுதிகள் (enumeration-blocks) வழிநிலைக் கூறுகளாகவும் (second stage units) கொள்ளப்பட்டு மாதிரிமுறை அமுலாக்கப்படுகிறது. பின்வரும் முறையில் படுகைகள் (strata) அமைக்கப் பெறுகின்றன.

முதற்படுகை :		100,000-ம் அதற்கு மேலும்	மக்கட்தொகை உள்ள நகரங்கள்
இரண்டாம்	,, :	50,000 முதல் 99,999 வரை	,, ,,
மூன்றும்	,, :	20,000 முதல் 49,999 வரை	,, ,,
நான்காம்	,, :	20,000 -க்குக் கீழ் மக்கட் தொகை	,, ,,

{ எல்லா மாநிலங்களிலும், முதற் படுகையில் ஒரே ஒரு நிலையில் மாதிரிமுறை கையாளப்படும்; இரண்டாம் படுகையில், சில மாநிலங்களைப் பொறுத்த அளவில், ஒருநிலை மாதிரி முறை கருதப்படுகிறது.}

ஒவ்வொரு மாநிலத்திற்கும் ஒதுக்கப்படும் தொகுதிகள் (blocks), மேற்கண்ட நான்கு படுகைகளிலும், அவற்றின் மக்கட் தொகையோடு வீதத் தொடர்பு உள்ள முறையில், பகிர்ந்து அளிக்கப்படுகின்றன. முதற்படுகையில், எல்லா நகர்களுமே தெரிவு செய்யப்படுகின்றன. இந் நகரங்களில், தொகுதிகள் அவற்றின் மக்கள்தொகை அளவு விகிதத்தில் பகிர்ந்து அளிக்கப் படுகின்றன.

இரண்டாவது படுகையில், படுகைகளில் பங்கிடப்படுக் மொத்தத் தொகுதிகளில் பாதியினைவிடக் குறைந்த அளவில் நகர்கள் எண்ணிக்கையானது அமைந்திருக்கும் நிலையில், முதலில் குறிப்பிட்ட அதே முறைப்படி பகிர்ந்தளித்தல் (distribution) நடைபெறுகிறது. தொகுதிகளை வெவ்வேறு நகர்களில் பங்கீடு செய்கையில், ஒவ்வொரு நகரும் குறைந்தபட்சமாக இரண்டு தொகுதிகளையேனும் பெறும்படியாகப் பார்த்துக்கொள்ளப்படுகிறது. மாதிரி நகர்களும் (Sample towns), அவற்றின் தொகுதிகளும் (blocks) எளிய சமவாய்ப்பு முறையின் அடிப்படையில் தெரிவு செய்யப்படுகின்றன. இது மாதிரி எடுப்பதற்கு மேற்கொள்ளப்படும் முறையாகும்.

பொதுவாக, ஒவ்வொரு மாநிலத்திலிருந்தும், 150 கிராமப்புறக் கூறுகளும், 100 நகர்புறத் தொகுதிகளும் தெரிவுசெய்யப்பட்டன. இதற்கான வழிமுறையினை மேலே குறிப்பிட்டோம். சிறிய அளவில் அமையும் மாநிலங்களில் (அசாம், மணிப்பூர், திரிபுரா போன்று), மாதிரி அளவுகளும் சிறியனவாக அமைந்தன, மொத்த இந்தியப் பரப்பிலிருந்தும், 2400 கிராமங்களும், 1300 நகர்ப்புறத் தொகுதிகளும், SRS திட்ட முறைக்குத் தெரிவுசெய்யப்பட்டன. இந்தத் திட்டத்தில் மொத்தம் 350 இலட்சம் மக்கள் இடம் பெற்றனர்.

கள ஆய்வு (Field work)

மற்ற மாதிரி அளவெடுப்புப்பணிகளைப்போலவே, SRS-க்குத் தேவைப்பட்ட பிறப்பு-இறப்பு நிகழ்வு விவரங்களைச் சேகரித்தற் பொருட்டு, பொதுநலப் பணியாளர்கள், ஆசிரியர்கள் போன்ற பகுதிநேரக் கணக்கெடுப்பாளர்கள் (part-time enumerators) பயன்படுத்தப் பட்டனர். இப்பணியில் அமர்த்தப்படுவதற்குமுன் அவர்கட்குத் தீவிரமான அளவில் செயல்பற்றிய போதுமான பயிற்சி அளிக்கப்பட்டது. கிராமப்புறப் பகுதிகளில், இக் கணக் கெடுப்பாளர்கள் தம்முடைய பெருமுயற்சியின்மூலமாகப் பல் வேறு பிரிவினரையும் நேரடியாகக் கண்டு, ஆய்வு நடத்தி, பிறப்பு-இறப்பு நிகழ்ச்சிகளைப்பற்றிய தெளிவான விவரங்களைப் பெற்று வருகின்றனர். இதன்பொருட்டு, இக் கணக்கெடுப்போர் மிக்க கவனத்துடன் தெரிந்தெடுக்கப்படும் தகவல் தருவோரது உதவியினை மிகவும் நாடுகின்றனர். உதாரணமாக, கிராமங்களில் உள்ள பல தொழில்புரிவோர், ஒன்றிய ஊழியர்கள், மத குரு மார்கள் போன்ற பலருடன் தொடர்பு கொண்டு பிறப்பு-இறப்பு நிகழ்ச்சிகளைப்பற்றிய விவரங்கள் திரட்டப்படுகின்றன. இதற்கு மாறான சூழ்நிலையில், நகர்ப்புறங்களில் நடத்தப்படும் ஆய்வில், கணக்கெடுப்போர் தாமே நேரடியாக எல்லாக் குடும்பங்களையும் மாதம் ஒருமுறையேனும் சென்று காண்பது தேவைப்படும் ஒன்றா கும். ஏனெனில், தகவல் அளிப்பவர்கள்மூலம் விவரங்கள் பெறு வது இந்நிலையில் நம்பத் தக்கதாய் அமைவதில்லை எனலாம்.

பிறப்பு, இறப்புபற்றிக் கண்டறிந்த விவரங்களைக் குறித்திடும் பொருட்டு, ஆய்வாளர்கட்குப் பொருத்தமான மனுத் தாள்கள் (forms) கொடுக்கப்படுகின்றன. ஒவ்வொரு மாத முதல் தேதியி லும், ஒவ்வொரு ஆய்வாளரும் தம் ஆய்வுபற்றிய விரிவான மாத அறிக்கையினை (சென்ற மாத அளவில்) தகுந்த வடிவில் தமது மேலதிகாரிகட்குச் சமர்ப்பிக்கவேண்டும். (தகுந்த நகல்களுடன்) அதன்பின், அந்த அறிக்கையானது, இந்தியப் பொதுப் பதிவாள ருக்கு (Registrar General) அனுப்பிவைக்கப்படுகிறது, இதுவே, SRS திட்டத்தின் ஆய்வுப்பணி பற்றிய விவரமாகும்.

அரை ஆண்டு அளவெடுப்பு (தனித் தன்மை உடையது) :—

முன்பு கருதப்பட்ட ஆய்வுடன் எந்தவிதத் தொடர்பும் இன்றித் தனியாக மேற்கொள்ளப்படும் ஆய்வு ஒவ்வொரு அரை ஆண்டுப் பகுதிக்கும் கருதப்பட்டு, அவற்றின்மூலம் சில துணை விவரங்கள் பெறப்பட்டு, முழு அளவினதான விவரங்கள் பதியப் படுகின்றன. ஒவ்வொரு கணக்கெடுப்பு ஆய்வாளரும், தமது அன்றாட ஆய்வுப் பணியில், எல்லா நிகழ்வுகளையும் (events

முழு அளவில் அறிந்து, விவரங்களைக் குறித்துவைப்பதை முக்கிய கடமையாகக் கருதுதல் இன்றியமையாதது. ஆனால், அவர்களை யும் அறியாமல் சில நிகழ்ச்சிகள் கணக்கெடுக்கப்படாமல் போக லாம். அல்லது சில விவரங்களைத் தவறாகக் குறிக்கவும் நேரிடும். எனவே, ஒவ்வொரு ஆறுமாத அளவிலும் (Six monthly) முழு நேரக் கண்காளிப்பாளர்மூலமாக, தனியான (independent) அள வில் ஆய்வு மேற்கொள்ளப்பட்டு, பிறப்பு-இறப்பு விவரங்கள் தனியாக, கூடவே சேகரிக்கப்படுகின்றன.

இதன்படி, ஆய்வு நடத்தும் கண்காணிப்பாளர் அல்லது மேற்பார்வையாளர், ஒவ்வொரு வீட்டுக்கும் நேரடியாகச் செல் வதன்மூலம், தெரிவு செய்யப்பட்ட மாதிரிக் கூறினைப் பொறுத்த அளவில், ஆறுமாத காலத்தில், சாதாரணமாக வசிக்கும் நபர் களிடம் மற்றும் வருகைதரும் நபர்களிடம் (Visitors) ஏற்படும் பிறப்பு மற்றும் இறப்புபற்றிய விவரங்களைத் திரட்டவேண்டும். இதோடு கூடவே, வீடுகள் மற்றும் குடும்பங்கள்பற்றிய விவரப் பட்டியல், மேலும் முதலில் ஆய்வாளர்கள் பூர்த்தி செய்த வினாப் பட்டியல் ஆகியவற்றில் தக்க மாற்றங்கள்செய்து, நிகழ்காலத் துக்கு ஏற்ற வகையில் மாற்றி அமைக்கப்படுகின்றன. ஏற் கெனவே நடந்த ஆய்வில் விட்டுப்போகும் அல்லது இடம்பெறாத விவரங்கள் இம்முறை சேர்க்கப்படவும் கூடும்.

இத்தகைய தனிப்பட்ட ஆய்வினை நடத்த உதவியாக, ஏற் கெனவே பெறப்பட்ட பிறப்பு-இறப்பு பற்றிய குறிப்பேடுகளைப் பயன்படுத்தல் இயலாது. ஏனெனில் அவை அனைத்தும் திரும்பப் பெறப்பட்டு, புதிய விவரங்கள் பெறப்படுகின்றன. பொதுவாக, ஒவ்வொரு ஆறு மாத காலத்திற்கும், கணக்கெடுப்பாளர்களால் புதிய பிறப்பு-இறப்புப் பற்றிய குறிப்பேடுகள் பயன்படுத்தப்படு கின்றன.

இதோடு நில்லாமல், மேற் குறிப்பிட்ட பகுதிநேரக் கண் காணிப்பாளர் (part-time supervisor) அவ்வப்போது (ஒரு கால அடிப்படையில்) மாதிரிக் கூறுகளைச் சென்று நேரடியாகக் காண் பதன்மூலமாக, ஆய்வாளர்களின் பணியினை மேற்பார்வையிட வும் செய்தல் வேண்டும். தேவையான விளக்கங்களையும், உதவி களையும் தருவதும் அவர்தம் கடமையாக அமைகிறது. இவ்வாறு சென்று காணும் தருணங்களில், மொத்தக் குடும்பங்களிலிருந்து (தெரிந்தெடுக்கப்பட்டவை), சுமாராக 10% அளவு கூறுகளை மட்டும் கருதி, அவற்றில் ஆய்வு நடத்துவதன் மூலமும், சம வாய்ப்பின் அடிப்படையில் சரிபார்க்கும் (check) முறை மூலமாக வும், பிறப்பு-இறப்புபற்றிய கணக்கெடுப்பில் பெறப்பட்ட விவரங்

களின் முழுமைத் தன்மையினையும், உறுதிப்பாட்டையும் சரிவர அறிதல் வேண்டும். இதில் ஏற்பட்ட குறைபாட்டினால், முழுநேரக் கண்காணிப்பாளர்கள் நியமிக்கப்பட்டு, களப் பணியின் மேற் பார்வை சிறந்து விளங்கவும், அரை ஆண்டு ஆய்வுகள் குறித்த காலத்தில் நடத்தி முடிக்கப்படவும் முயற்சிகள் எடுக்கப் பட்டுள்ளன.

ஒப்பீடு (Matching the Results)

அரை ஆண்டு ஆய்வு முற்றுப்பெற்றபின்னர், அதன்மூலம் பெற்ற முடிவுகள் அனைத்தும், ஆய்வாளர்களால் குறிக்கப்பட்டு மேலிடத்திற்கு அனுப்பப்பட்ட பிறப்பு-இறப்பு பற்றிய விவரக் குறிப்புகளுடன் ஒப்பிடப்படுகின்றன. ஆய்வாளரின் குறிப்பேட்டி லும், கண்காணிப்பாளரின் குறிப்பேட்டிலும் இடம்பெறும் குறிப்புகள் (entries) அனைத்தும், ஒவ்வொரு உறுப்பாகக் கருதி இணைக்கப்பட்டு, சரிபார்த்து ஒப்பிடப்படுகின்றன. நிகழ்வுகள் (events) அனைத்தும், 'முற்றிலும் ஒத்திருப்பவை' என்றும், 'பகுதிமட்டும் ஒத்திருப்பவை' (partly matched) என்றும், 'ஒத்திராதவை' (unmatched) என்றுமாகப் பிரிக்கப்படுகின்றன. இத்தகைய 'இணைத்து ஒப்பிடுதல்' மூலமாக, விட்டுப்போன விவரங்களை (omission) கண்டறியவும், விவரங்களின் சரியான தன்மையினை (correctness) அறிந்திடவும், தவறாகக் குறிப்பிட்ட விவரங்களை இனம் காணவும் வழி ஏற்படுகின்றது. எனவே, இரு வேறு விதங்களில் பெற்ற விவரங்களின் ஒப்புமையானது, சரியான பிறப்பு-இறப்பு நிகழ்வுபற்றிய விவரங்களைப் பெறுவதில் மிக்க உதவி புரிகிறது.

'இணைத்து ஒப்பிடுதலில்' (matching), பொதுவாகப் பின் வரும் உண்மைகள் கருதப்படுகின்றன :

1. குடும்பு அமைப்பு : அதாவது, குடும்பத் தலைவரின் பெயர், மற்றும் வீட்டு எண்.
2. தாயாரின் பெயர் (பிறப்புகளைக் கண்டறிய உதவுவது), மற்றும் இறந்தவர் பெயர் (இறப்புகளைப் பொறுத்து).
3. வாழும் நிலை (இயல்பாக வசிப்பவரா ; விருந்தினரா).
4. இனம் (ஆண் பெண்).
5. நிகழ்வு நடைபெற்ற மாதம்
6. இறக்கும் தறுவாயில், (இறந்தவரின்) வயது.

இந் நிலையில், ஒப்பிடுதலானது முதல் நிலை 'இணைத்து ஒப் பிடுதல்' என்று அறியப்படும். ஆய்வுக் களத்திலேயே அடுத்த நிலை (அல்லது கட்ட) ஒப்பிடுதல் அமைவதை இனி காண்போம்.

ஆய்வுக்கள் ஒப்பீடு (Field Verification)

இரு ஆய்வுகளிலும் விவரங்கள் முற்றிலும் ஒத்திருப்பின். எந்தவிதப் பிரச்சினையும் வருவதில்லை. எனவே, 'பகுதி மட்டும் ஒத்திருக்கும்' (partly matched), மற்றும் 'ஒப்புமையில்லாத' (unmatched) நிகழ்வுகளைத் திரும்பவும் ஆய்வுக் களத்திற்கே செல்வதன் மூலமாகத் திருத்தவேண்டிய நிலை ஏற்படுகிறது, இவ்விவரங்களின் சரியான நிலையானது களத்தில் அறியப்படுகிறது. இத்தகைய மறு ஒப்பீட்டுப் பணியானது, மூன்றுவது மனிதர் மூலமாக, அல்லது கண்காணிப்பாளர் மற்றும் ஆய்வாளர் இருவரும் இணைந்து பணியாற்றுதல்மூலமாக, நடைபெறுகிறது. கிடைக்கும் பணியாளரைப் பொறுத்து இது அமைகிறது. பெரும்பாலும், தவறான விவரங்கள் திருத்தப்பட்டும், மாற்றப்பட்டும் அமையும் நிலையானது ஏற்படாவண்ணம், மூன்றுவது நபர் மூலமாகவே இத்தகைய ஒப்பிடுதல் நடத்தப்படுகிறது. மேலும், அரைவருட ஆய்வுக் காலத்திலும், திரும்ப ஒப்பிடும் காலத்திலும் நடைபெறும் சில புதிய நிகழ்வுகள் பற்றிய விவரங்களும் சேகரிக்கப்படுகின்றன.

ஒவ்வொரு அரை ஆண்டு அளவெடுப்பும் முடிவுபெற்ற நிலையில், இணைத்து ஒப்பிடுதல், மற்றும் மறுமுறை ஒப்பிடுதல் ஆகியவற்றின்மூலம் பெறப்படும் இறுதி முடிவுகளும், ஒவ்வொரு கூறினுக்கும், இனம் மற்றும் வயது அடிப்படையில் மக்கட் தொகை பற்றிய விவரங்களும், தகுந்த படிவங்களில் (form) மாநில மேலிடங்கள்மூலம் பொதுப் பதிவாளருக்கு அனுப்பி வைக்கப்படுகின்றன. அவற்றின் வாயிலாக, அட்டவணைப் படுத்துதல், விவரங்களின் பகுப்பாய்வு ஆகிய செயல்முறைகள் தீவிரமாக இயங்கத் துவங்குகின்றன.

இதுவரை கருதப்பட்ட பகுதி, *SRS* திட்ட முறையின் செயல்முறை பற்றிய விரிவான ஆய்வாக அமைந்தது. இனி வரும் பகுதியில், கணக்கீட்டு அடிப்படையில், வெவ்வேறு மதிப்பீடுகள் பெறப்படும் வழிமுறையானது விரிவான அளவில் கருதப்பட்டு, ஆராயப்படுகிறது.

மதிப்பீட்டு முறை (நகர்ப்புற அளவில்) (Estimation Method—Urban Sector)

குறியீடு :

N_h : h-ஆவது படுகையில் மொத்தம் உள்ள நகர்கள்.

n_h : h-ஆவது படுகையில் மாதிரி நகர்கள்.

M_{hi} : h-ஆவது படுகையில், i-ஆவது நகரில் உள்ள தொகுதிகள்.

m_{hi} : ,, ,, மாதிரித் தொகுதிகள்.

Y_{hij} : h-வது படுகை, i-வது நகர், j-ஆவது தொகுதியில் மொத்த நபர்கள்.

b_{hij} : h-வது படுகை, i-ஆவது நகர் j-வது தொகுதியில் மொத்தப் பிறப்புகள்.

மொத்தம், மற்றும் வீதங்கட்கு மதிப்பீடு.

h-வது படுகையின் மொத்த மக்கள்தொகையின் பிறழ்ச்சியற்ற மதிப்பீடாக ($h = 1, 2, 3, 4$) $\hat{Y}_h$ என்பதையும், மொத்த மாநிலத்திற்கும் இதனை $\hat{Y}$ என்றும் குறிப்பிடுக.

முதல் படுகையில், ஒரே நிலை கருதப்படுவதாலும், ஏனைய வற்றில் இருநிலைகள் கருதப்படுவதாலும்,

$$\hat{Y} = \hat{Y}_1 + \hat{Y}_2 + \hat{Y}_3 + \hat{Y}_4$$

என்று அமைகிறது.

இதில்,

$$\hat{Y}_1 = \sum_{i=1}^{N_1} \frac{M_{1i}}{m_{1i}} \cdot \sum_{j} Y_{1ij} = \sum_{i} M_{1i} \cdot \bar{Y}_{1i}$$

என்றும்,

$$\hat{Y}_h = \frac{N_h}{n_h} \sum_{i=1}^{n_h} \frac{M_{hi}}{m_{hi}} \sum_{j=1}^{M_{hi}} Y_{hij}$$

$$= \frac{N_h}{n_h} \sum_{i} M_{hi} \cdot \bar{Y}_{hi} \quad (h = 2, 3, 4)$$

எனவும், இருவிதமாக அமையும்.

இதேபோல், மாநிலத்தில் மொத்தமாக உள்ள பிறப்புகளின் எண்ணிக்கைக்கும் ஒரு பிறழ்ச்சியற்ற மதிப்பீடு பெறப்படும்.

மாநில முழுதுமாக, உள்ள பிறப்பு வீதமானது

$$\frac{\hat{B}}{\hat{Y}} \times 1000 = r \times 1000 \text{ (என்று கொள்க)}$$

என்று மதிப்பிடப்படும். ($\hat{B}$ பிறப்பு அளவின் மதிப்பீடு)

மாதிரிமுறை சார்ந்த மாறுபாட்டின் மதிப்பீடு: (Sampling Variance)

$V(\hat{Y})$-ன் பிறழ்ச்சியற்ற மதிப்பீடு

$$V(\hat{Y}) = V(\hat{Y}_1) + V(\hat{Y}_2) + V(\hat{Y}_3) + V(\hat{Y}_4)$$

என்று தரப்படும்.

இதில்,

$$V(\hat{Y}_1) = \sum_{i=1}^{N_1} \frac{M_{1i}}{m_{1i}} \frac{(M_{1i} - m_{1i})}{(m_{1i} - 1)} \cdot \sum_{j=1}^{m_{1i}} (Y_{1ij} - \bar{Y}_{1i})^2$$

என்றும்,

$$V(\hat{Y}_h) = \frac{N_h (N_h - n_h)}{n_h (n_h - 1)} \sum_{i=1}^{n_h} (M_{hi} \cdot \bar{Y}_{hi} - Y_h / N_h)^2$$

$$+ \frac{N_h}{n_h} \sum_{i=1}^{n_h} \frac{M_{hi} (M_{hi} - m_{hi})}{m_{hi} (m_{hi} - 1)} \cdot \sum_{j=1}^{m_{hi}} (Y_{hj} - \bar{Y}_{hi})^2$$

என்றும் பெறப்படும். ($h = 2, 3, 4$)

இதேபோல, $V(\hat{B})$-ன் மதிப்பீடும் பெறப்படுகிறது.

$\hat{B}$, மற்றும் $\hat{Y}$ இவற்றின் இடையே உள்ள உடன் மாறுபாடும் பின்வருமாறு பிறழ்ச்சியற்று மதிப்பிடப்படும்.

$$Cov\ (B, \hat{Y}) = \sum_{h=1}^{4} Cov\ (\hat{B}_h, \hat{Y}_h).$$

இதில்,

$$Cov(\hat{B}_1, \hat{Y}_1) = \sum_{i=1}^{N_1} \frac{M_{1i}(M_{1i} - m_{1i})}{m_{1i}(m_{1i} - 1)} \sum_{j=1}^{m_{1i}} (Y_{1ij} - \bar{Y}_{1i})(b_{1ij} - b_{1i})$$

என்று அமையும்.

$Cov(\hat{B}_h, \hat{Y}_h)$-ன் மதிப்பையும் $V(\hat{Y}_h)$ போல் பெறலாம்.

பிறப்பு வீத மதிப்பீட்டின் மாறுபாட்டு அளவு

$V(B.R) = (1000)^2.\ V(r)$ என்றாகும்.

[$B.R$: பிறப்பு வீதம் (Birth Rate)]

இதில்,

$$V(r) = \frac{1}{(\hat{Y})^2}.\ [\,V(\hat{B}) + r^2.V(\hat{Y}) - 2r \cdot Cov(\hat{B}, \hat{Y})\,]$$

என வரும்.

BR-ன் மாறுபாட்டுக் கெழு (Co-efficient of Variation).

$$e = \frac{\sqrt{V(r)}}{r} \times 100$$ என வரும்.

கிராமப்புற மதிப்பீடுகள் (Estimation—Village Sector)

N_h : h-வது படுகையில் உள்ள மொத்த கிராமங்கள் அல்லது பிரிவுகள். (segments)

n_h : h-வது படுகையில் மாதிரி ,, ,,

Y_{hi} : h-வது படுகையில், i-ஆவது மாதிரி கிராமத்தில் உள்ள நபர்கள்.

b_i h-வது படுகை, i-வது கிராமத்தில் ,, ,,

மதிப்பீடுகள் :

மொத்த மதிப்பீடு :

$\hat{Y}$ என்பது மாநில மொத்த மக்கட் தொகையின் பிறழ்ச்சியற்ற மதிப்பீடெனில்

$$\hat{Y} = \sum_h \frac{N_h}{n_h} \sum_i Y_{hi}$$

$$\therefore \quad \hat{Y} = \sum_{1}^{L} N_h \cdot \bar{Y}_h \quad \text{இதில் } \bar{Y}_h = \sum_{1}^{n_h} \frac{Y_{hi}}{n_h}.$$

இதேபோல், $\hat{B}$ மதிப்பீடானது

$$\hat{B} = \Sigma \; N_h \cdot \bar{b}_h \qquad \left(\bar{b}_h = \sum_{i} \frac{b_{hi}}{n_h} \right)$$

என்று வருகின்றது.

$\therefore$ பிறப்பு வீதத்தின் மதிப்பீடு :

$\frac{\hat{B}}{\hat{Y}} \times 1000 = r \times 1000$ என்று பெறப்படும்.

மாறுபாடு (Varlance).

$V(\hat{Y})$-ன் பிறழ்ச்சியற்ற மாறுபாடு.

$$V(\hat{Y}) = \sum_{h} \frac{N_h (N_h - n_h)}{n_h (n_h - 1)} \sum_{1}^{n_h} (Y_{hi} - \bar{Y}_h)^2$$

என்றும், இதேபோல் $V(\hat{B})$-ம் கொடுக்கப்படும்.

$$Cov(\hat{Y}, B) = \sum_{h} \frac{N_h (N_h - n_h)}{n_h (n_h - 1)} \cdot \sum_{i} (Y_{hi} - \bar{Y}_h)(b_{hi} - \bar{b}_h)$$

என்பது இவற்றினிடையே உடன் மாறுபாடாகும் (Covarince)

$\therefore$ பிறப்பு வீதத்தின் மாறுபாடு.

$$V(B.R) = (1000)^2 \times V(r) = (1000)^2 \cdot \frac{1}{\hat{Y}^2}$$

$$[\, V(B) + r \cdot V(Y) - 2r \cdot Cov(\hat{B}, \hat{Y}) \,]$$

என்றாகும்.

III இந்திய விவசாயப் புள்ளி விவரங்கள் (Indian Agricultural Statistics)

இந்தியப் பொருளாதாரத்தின் முதுகெலும்பாக விளங்கும் விவசாயத் துறையில், கிராப்புறச் சூழ்நிலைபற்றிய அளவில், பல்வேறு இயல்புகளையும்பற்றி விரிவான ஆய்வுகள்மூலம் விவரங்கள் (Statistics) சேகரிக்கப்படுகின்றன. கால்நடை நிலை, வனவளம் (Forestry) மற்றும் விவசாயத் தோடு மிகவும் தொடர்புடைய தொழில்களில் நிலை ஆகிய பலவும் சேர்ந்து ஏறக்குறைய இந்திய தேசீய வருமானத்தில் (National Income) பாதியைத் தீர்மானிக்கின்றன எனலாம். எனவே, நாட்டின் வெவ்வேறு பகுதிகளில் ஏற்படும் பஞ்சம், வெள்ளம் பருவமழை பெய்யாத நிலை போன்றவை எந்த விதத்தில் நாட்டின் விவசாயத் துறையைப் பாதிக்கின்றன என்று நன்கறியவும், தகுந்த முன்னெச்சரிக்கைகளை மேற்கொண்டு விவசாய உற்பத்தி குறையாது இருக்கச் செய்யவும் உதவத்தக்க அளவில் இவ் விவரங்கள் அதிக நம்பகத்தன்மையுடனும், போதிய அளவிலும் திரட்டப்பட வேண்டியிருக்கிறது. ஒவ்வொரு பயிரும் பயன்படுத்தப்பட வேண்டிய நிலப்பரப்பும் (acreage), கிடைக்கும் உற்பத்திப் பொருளின் அளவிற்கேற்பத் தொழில் உற்பத்தியின் அளவும், இவற்றின் மூலம்தான் மதிப்பிடப்படுகின்றன. விவசாய அபிவிருத்திக்கு நிதி உதவி செய்ய வங்கிகளுக்கு முழு அளவில் முக்கிய பயிர்கள்பற்றிய விவரங்கள் தேவைப்படுகின்றன.

விவசாயப் புள்ளி விவரங்கள், முக்கியத் தன்மையுடன்

(I) நிலப்பரப்பு அல்லது நிலப் பயன்பாட்டு விவரங்கள் (Area or Land Utilisation Statistics)

(ii) உற்பத்தி அல்லது விளைச்சல் புள்ளவிவரங்கள் (Yield Statistics)

என இருவகையாகப் பிரிக்கப்படுகின்றன.

நிலப் பயன்பாட்டு விவரங்கள் (Land Utilisation)

நாட்டில் பூகோள எல்லை வரையறையில் இடம்பெறும் நிலப்பரப்பு அனைத்தும் எவ்வாறு பயன்படுகிறது என்பதைப் பற்றிப் பெறப்படும் விவரங்கள் அனைத்தும், விவசாயத்தைப் பொறுத்த அளவில், இப் பிரிவில் இடம் பெறுகின்றன. இதன் பொருட்டு, மொத்த நிலப்பரப்பு, வனத்தில் அடங்கும் பரப்பு, விதைக்கப்படும் நில அளவு (area sown), வெவ்வேறு பயிர்கள்

பயனாகும் நிலப்பரப்பு, பாசனம் செய்யப்படும் பயிர்கள் (Irrigated) போன்ற பல விவரங்களும் சேகரிக்கப் படுகின்றன. ஆண்டு தோறும் வெளியாகும் (Indian Agricultural Statistics) என்னிம். மத்திய உணவு, விவசாயம், சமூக முன்னேற்றம் பற்றிய அமைச்சகத்தின் வெளியீட்டில் இவ் விவரங்கள் அனைத்தும் தெளிவாகத் தரப்படுகின்றன. இதன்மூலம், (அ) மொத்தப் பரப்பு ; (ஆ) நிலப்பரப்பின் பாகுபாடு (9 பிரிவுகளாக) ; (இ) பாசனம் செய்யப்படும் பகுதியும் பயிர்களும் ; (ஈ) பயிர்கள் பயன்படுத்தப்படும் நிலப்பரப்பு போன்ற பல்வேறு விவரங்கள் தொகுத்தளிக்கப்படுகின்றன.

நிலப்பரப்பு மற்றும் சாகுபடி (Yield) பற்றிய அளவில் இங்கு சிலவற்றைக் குறிப்பிடுவோம்:

(1) **நிலப்பரப்பு (Area)**

இதில், வெவ்வேறு பயிர்களை விளைவிக்கக் கருதப்படும் பரப்பும், மற்ற வகைகளில் பயன்படுத்தப்படும் நிலப்பரப்பும் இடம் பெறுகின்றன. நாடு முழுதும் மொத்தமாக எடுக்கப்பட்டு, இரண்டு பிரிவுகளாக அமைக்கப்பட்டுள்ளன. அவையாவன :

(i) **தற்காலிகத் தீர்வு காணப்பட்ட பரப்புகள்** (Temporarily-settled Area)

(ii) **நிரந்தரத் தீர்வு காணப்பட்ட பரப்புகள்** (Permanently-settled Areas)

(i) தற்காலிகத் தீர்வு காணப்பட்ட பகுதிகளான தமிழ்நாடு, பஞ்சாப், உ.பி. போன்ற மாநிலங்களில் பின்வருமாறு காணலாம்: ஒவ்வொரு கிராமத்திலும் அதன் கணக்கு வழக்கு விவகாரங்கட்குப் பொறுப்பான கணக்காயர் அல்லது கர்ணம் (Village Accountant) மூலமாக, பயிர் செய்யப்படும் நிலப்பரப்புகள்பற்றிய எல்லா விவரங்களும் சேகரிக்கப்பட்டு பெறப்படுகின்றன. இதற்கென அமைந்துள்ள குறிப்பேட்டில் அல்லது பதிவேட்டில் (register) விவரங்கள் சரியாகக் குறித்து வைக்கப்படுகின்றன. நிலங்களுக்குத் தக்கபடி ஆய்வு எண்கள் கொடுக்கப்பட்டு (survey numbers), நிலத்தின் அளவு, அதன் மதிப்பு, உடைமை (ownership) போன்ற விவரங்கள் ஒவ்வொரு நிலப் பகுதிக்கும் பெறப்படுகின்றன கணக்காயரது பணிகள் வருவாய் ஆய்வாளரின் (Revenue Inspector) மூலம் மேற்பார்வையிடப்

படுகின்றன. மேல் மட்டத்தில், தாலுகா உயர் அதிகாரிகளால் (Taluk Officials) விவரங்கள் சரிபார்க்கப்படும். ஒவ்வொரு மாவட்டத்திலும், வெவ்வேறு பயிர்கள் விளைவிக்கப்படும் மொத்த நிலப்பரப்பு (total area) மதிப்பிடப்படும். இறுதியில் மாநில மதிப்பீடு கணக்கிடப்படும்.

(ii) நிரந்தரத் தீர்வு காணப்பட்ட இடங்களான பீகார், ஒரிசா, வங்காளம் போன்ற பகுதிகளில், நிரந்தர அமைப்பாக முக்கிய அறிவிப்பாளர்கள் (reporters) அமைந்திருப்பதில்லை. எனவே, விவரங்களும் திருப்திகரமான அளவில் அமைவதில்லை. கிராமத் தலையாரி (headman) அல்லது வேறு ஏதாவது அதிகாரிமூலமாக, நிலவரி மற்றும் அதனோடு தொடர்புடையனவாக உள்ள வருவாய் (land revenue) பற்றிய விவரங்கள் பெறப்படும். இவ்விவரங்கள் தக்கபடி சரிபார்க்கப் படுவதில்லை. எனவே, நனிநபர் மதிப்பீடுமூலம் விவரங்கள் அமைகின்ற இந் நிலையில், விவரங்கள் தவறுனவைகளாக இருப்பது குறிப்பிடத்தக்கது. பொதுவாக, இறுதியில் விவசாய இயக்குநர் (Director of Agriculture) மூலமாக மதிப்பிட்டு வெளியிடப்படும் மொத்த நிலப்பரப்பு தவறுடையதாக அமைகிறது

எனவே, இந்நிலையில், வெவ்வேறு பயிர்களின் நிலப்பரபினைக் குறித்துவைக்க, தக்கபடி திட்டமிட்ட அமைப்பு உருவாக்கப் படல் இன்றியமையாதது. எல்லா மாநிலங்களிலும் ஒரே சீரான (uniform) வழிமுறை பின்பற்றப்படல் வேண்டும். ஒவ்வொரு தாலுக்கா அல்லது தாசில் ஆகியவற்றின் வரைபடங்களைக் கொண்டு, முறையான ஆய்வு நடத்தப்படவேண்டும். முழு அளவில் கணிப்பு ஆய்வு நடத்துதல் அவசியமாகிறது. இதற்கேற்ற நடவடிக்கைகள், தற்கால அளவில் துரிதமாக எடுக்கப்பட்டு ஓரளவு வெற்றியும் பெறப்பட்டு வருகின்றது.

(2) விளைவு அல்லது உற்பத்தி விவரங்கள் (Yield Statistics)

இதன்மூலம் பல்வேறு உணவுப் பயிர்கள் மற்றும் இதரப் பயிர்கள் பற்றிய பயிர்ச் சாகுபடி அல்லது உற்பத்தி அளவுகள் மதிப்பிடப்படுவதைக் குறிப்பிடுகிறோம். அதாவது, பயிர் செய்ய அமையும் நிலப்பரப்பில் விளையும் பொருள்களின் மொத்த உற்பத்தி மதிப்பீட்டினை இங்கு எடுத்துக்கொள்கிறோம். இதரப் பயிர்களாக (nonfood crops), பருத்தி, எண்ணெய்வித்துக்கள். போன்றவை அறியப்படுகின்றன. மொத்தவிளைவு (yield) பற்றி இருவகையான விவரங்கள் வெளியாகின்றன.

(1) அரசினரின் வெளியீட்டு வரிசைகள் (official series) இவை பெரும் அளவில் பயிர்களைக் கருத்திற்கொண்டு பெறப் படும் விரிவான விவரங்கள் அமைந்திருப்பவையாகும். இருவேறு வழிகளில், பயிர் விளைச்சல்கள் அல்லது பயிர்சாகுபடிகள்பற்றிய விவரங்கள் நம் நாட்டில் திரட்டப்படுகின்றன. அதாவது (அ) வழக்கமான முறை (Traditional Method) ; (ஆ) வாய்ப்பின் அடிப்படையில் அமையும் மாதிரிமுறை (Random Sampling Method).

வழக்கமான (traditional) முறை என்பது,

மொத்த உற்பத்தி = பயிரின் விளைச்சல் பரப்பு × இயல்பான உற்பத்தி × நிபந்தனைக் காரணி (conditional factor)

'இயல்பான உற்பத்தியாவது 'சராசரியான மண்வளத்தில், ஓராண்டில், இயல்பான (சராசரியான) நிலையில், பெறப்படும் சராசரி உற்பத்தி (average out-turn)' என்பதாக ஒவ்வொரு பயிருக்கும் வரையறுக்கப்படுகிறது. இது மிகச் 'சிறந்த விளைச்சலை'யோ அல்லது 'அமோக விளைச்சலை'யோ (bumper crop) குறிப்பதல்ல. மாறாக, கருதப்படும் நிலப்பரப்பினைக் கொண்டு மதிப்பிடப்படும், குறிப்பிட்ட பயிரின் (typical crop) சராசரி விளைச்சல் ஆகும். முந்தைய அனுபவத்தின் அடிப் படையில் அமைந்திருக்கும். தொடர்ச்சியாகப் பல ஆண்டுகளைக் கருதிய அளவில் அறியப்படும். திரும்பத்திரும்ப அதிகஅளவில் பயிரிடப்படும் பொதுவான பயிரானது 'குறிப்பிட்ட பயிராக' அமையும். உதாரணமாக, ஒரு நிலப்பகுதியில் திரும்பத்திரும்ப சோளமானது அதிகமுறை பயிரிடப்படின், அதுவே அந் நிலத்தின் 'குறிப்பிட்ட பயிராக' அமைகிறது.

ஒவ்வொரு மாவட்டத்திலும், வெவ்வேறு பயிர்களுக்கு, ஒரு ஏகரில் பயிராகும் இயல்பான விளைச்சல் அளவினை நிர்ணயிப்பது விவசாயத் துறையின் (Dept. of Agriculture) பொறுப்பாகும். இதற்காக, சில நிலக்கூறுகள் (plots) பிரதிநிதிகளாகத் தெரிந் தெடுக்கப்படுகின்றன. இவ்வாறு தெரிந்தெடுக்கப்படும் நிலப் பகுதிகளில், பாசனம், எருவிடல் (manure) விதைவிதைத்தல் போன்ற பணிகளும் (அறுவடை முதற்கொண்டு) விவசாயத் துறை அதிகாரிகளின் மேற்பார்வையில் நடைபெறுகின்றன. மொத்த விளைச்சலானது எடையிடப்பட்டு முடிவுபெறப்படும்.

பயிர் விளைச்சல் (crop cutting) சோதனைகளில் பெறப்படும் விவரங்கள், இம் மதிப்பீடுகளை, (பயிர்விளைச்சல் பற்றிய)

பெறுகையில் நிலவும் சூழ்நிலைகளை நன்கு மனத்திற்கொண்டு அமைதல் வேண்டும். எனவே நிபந்தனைக் காரணி யானது' (அல்லது பருவகாலம் பற்றிய காரணியானது (Seasonal factor) இயல்பான விளைச்சலோடு தொடர்புடைய குறிப்பிட்ட பருவ காலத்தினைக் குறிப்பதாக உள்ளது. இத்தகைய நிபந்தனைக் காரணியானது பயிரின் வளர்ச்சிக் காலத்திலும், அறுவடைக் காலத்திலும் என்பதாக இரு வேறு காலங்களில் மதிப்பிடப் படுகிறது. லேக்பால் (Lekhpal) எனப்படும் அதிகாரிகள் தம் பகுதி விவரங்களை, தாசில்தார்கட்கு அளிக்க, அவர்மூலம் விவரம் சரிபார்க்கப்பட்டு, தாசில் அளவில் மதிப்பீடு பெறப்பட்டு, மாவட்ட அதிகாரிக்கு அனுப்பப்படும் இறுதியில் தக்க மாற்றங் களுடனும், திருத்தங்களுடனும் பெறப்படும் மதிப்பீடுகள் கருதப் பட்டு, மாநில அளவில் மதிப்பீடு பெறப்படும். இதில் அமைந் துள்ள பல்வேறு குறைபாடுகள் வருமாறு :-

'இயல்பான விளைச்சல்' (உற்பத்தி) என்பது தெளிவாக விளக்கப்படாத கருத்தாகும். அதனை மதிப்பிடுவது எளிதான செயல் அல்ல. தெரிந்தெடுக்கப்படும் நிலப்பகுதிகள் இருக்கும் சூழ்நிலையைப் பிரதிபலிப்பதாக அமையாமல் போகின்றன. இத்தகைய நிலப்பகுதிகட்குக் கிடைக்கும் பாசனம், விதைகள், எருக்கள் (உரங்கள்) போன்ற பலவசதிகள் இதர நிலங்களுக்குக் கிடைக்காமல் போகலாம். எனவே ஒரு நிலப்பரப்பின் உண்மை நிலையினைக் காட்டுவதாக, 'இயல்பான விளைச்சல்' மதிப்பீடு அமைவது இல்லை. மாறுபடும் சூழ்நிலைகளுக்கு ஏற்ற அளவில் (5 ஆண்டுகட்கு) இயல்பான விளைச்சல் மாற்றப்படுவதில்லை. எனவே, ஒரே மதிப்பைப் பயன்படுத்துவது (நீண்ட காலத்திற்கு) சரியல்ல. நிபந்தனைக் காரணி நிர்ணயிக்கப்படும் வீதமும் குறை பாட்டுடன் அமைந்திருந்க்கிறது

தக்க மாற்றத்துடன் 'இயல்பான விளைச்சல்' என்பது, 'பத்து ஆண்டுகளில் நடத்தப்படும் பயிர் விளைச்சல் ஆய்வு முடிவுகளின் அடிப்படையில் தீர்மானிக்கப்படும் உண்மையான விளைச்சல் அளவுகளின் (ஒரு ஏக்கரின்) நகரும் சராசரி (moving average)' என வரையறுக்கப்பட்டுள்ளது.

ராண்டம் அடிப்படையில் அமைந்த மாதிரி (Random Sampling)

ஊக மதிப்பீட்டிலும், பயிர்பற்றிய மதிப்பீட்டிலும் உள்ள குறைபாடுகளைக் களைதற்பொருட்டு, சமவாய்ப்பு அடிப்படையில் நிலக்கூறுகள் அமைந்த மாதிரியினைத் தேர்ந்து எடுத்து, பயிர் விளைச்சலானது மதிப்பிடும் நிலை மேற்கொள்ளப்படுகிறது. சில

மாநில அரசுகள் இம்முறையினை மேற்கொள்ளக் காணலாம். (உதாரணம்: பீகார், ஒரிசா—1923.) செலவோடு ஒப்பிடப் படுகையில், இதன்மூலம் பெறப்பட்ட முடிவுகள் நம்பிக்கை யூட்டத்தக்க விகிதத்தில் அமையவில்லை. 1942-ல், பருத்தி விளைவிக்கப்படும் நிலப்பரப்பில் பயன்படுத்தத்தக்க சமவாய்ப்பு மாதிரி ஆய்வு முறையினை 'இந்திய விவசாய ஆராய்ச்சிக் கழக' மானது (Indian Council of Agricultural Research)வழிவகுத்தது. இது சிறந்ததாகக் கருதப்பட்டு, உணவுப் பயிர்களின் விளைச்சலை மதிப்பிட, பஞ்சாப், தமிழ்நாடு, பம்பாய், மத்தியபிரதேசம்போன்ற பல மாநிலங்களில் பயன்படுத்தப்பட்டது. எல்லா மாநிலங் களிலும் இம்முறையானது (1963 முதல்) சமவாய்ப்பு மாதிரிமுறை மூலம் பயிர் விளைச்சல் ஆய்வுகளில் பயன்படுத்தப்பட்டு வருகிறது. **மாதிரி ஆய்வு** முறைகள்மூலம், விவசாய உற்பத்தி யினை மதிப்பிட்டறியப் பெரும் முயற்சிகள் மேற்கொள்ளப் படுகின்றன.

I. C. A. R. மூலம் கையாளப்படும் வழி முறையானது, படுகையாக்கப்பட்ட பலகட்ட ராண்டம் மாதிரிமுறை (Stratified Multistage Random Sampling) என அறியப்படுகிறது.

வழிமுறை: சமவாய்ப்பு மாதிரிமுறை விதிப்படி, ஒவ்வொரு தாலுகா விலும் (அல்லது தாசில்) பெரும் அளவில் கிராமங்கள் தேர்ந்தெடுக்கப்படுகின்றன. ஒவ்வொரு கிராமத்திலும் இரண்டு அல்லது மூன்று நிலக் கூறுகள் ராண்டமாக (at random) தேர்வு செய்யப்படுகின்றன. பரிசோதனைக்கு எடுக்கப்பட்ட நிலக்கூறு களின் அளவும் (size), வடிவமும் (shape) முடிவு செய்து கொள்ளப்படுகின்றன. தகுந்த பிரிவுகள் (demarcation) உருவாக்கப்படுகின்றன. பரிசோதனைக் கூறிலிருந்து அறுவடை செய்யப்பட்டு, அடித்துப் பதர்நீக்கிய நிலையில் (winnowed) எடையிடப்படும் உற்பத்தி அளவு பெறப்படுகிறது. இத்தகைய ஆய்வுச் செயல் முறைகளை I. C. A. R. வகுத்தளிக்கிறது.

இம் முறையானது அறிவியல் அடிப்படையில் சிறந்து அமைகிறது. முடிவுகளும், ஏற்கெனவே ஆராயப்பட்ட 'அன்னவாரி' அல்லது 'வழக்கமான' முறையைவிடச் சிறந்தவை யாகக் காணப்படுகின்றன. ஒரு ஏக்கருக்கு சராசரியாகப் பெறப் படும் உற்பத்தி அளவின் பிறழ்ச்சியற்ற மதிப்பீடு பெறப்படும் ஆனால், செலவு அதிகமாகவும், கடினமான ஆய்வுமுறைகளையும் கொண்டு அமைகிறது.

(ii) NSS அளவெடுப்புத் தொடர்கள் (NSS Series)

ஒழுங்கான ஆய்வுகளில், NSS மூலம், முக்கிய உணவுப் பயிர்களின் (Cereals) விளைச்சல் அல்லது உற்பத்திபற்றிய விவரங்கள் சேகரிக்கப்படுகின்றன. கிராமங்களில், இதர விவரங்களைச் சேகரிக்க அனுப்பப்படும் ஆய்வாளர்கள், பயிர் உற்பத்தி பற்றிய விவரங்களையும் திரட்டுகின்றனர். ஆனால், தற்காலிகமாக இடம்விட்டு இடம் செல்லும் NSS ஆய்வாவாளர் மூலமாக விவரங்களைப் பெறுவது அவ்வளவாகப் பயன்தருவதாக அமைவதில்லை. இருவேறு முறைகளிலும் பெரிதும் வேறுபாடுகள் காணப்படுகின்றன.

விவசாய உற்பத்திக் குறியீடுகள் (Indices of Agricultural Production).

நாட்டின் விவசாய உற்பத்தியின் அதிகரிப்பு வீதத்தைக் குறிப்பிட்டுக் காட்டுதற்கு, சில குறியீட்டெண்கள் கணக்கிடப்பட்டு அவ்வப்போது வெளியிடப்படுகின்றன. அவற்றுள் சில :

(i) உணவு, விவசாய அமைச்சகத்தினின்று வெளியிடப்படும் 'விவசாய உற்பத்திக் குறியீட்டெண்கள்' உணவு மற்றும் இதரப் பயிர்களைப் (28 முக்கியமானவை) பொறுத்த அளவில் கணக்கிடப்பட்டு, அவ்வப்போது தொடர்ச்சியான காலவரையில் வெளியிடப்படுகின்றன. சிறிய பயிர்கள் மற்றும் கால்நடை பற்றிய விவரங்கள் இதில் அடங்குவதில்லை. 1950 என்ற ஆண்டானது (ஜூன் மாதம்) 'அடிப்படை' (Base) ஆண்டாகக் கொள்ளப்படுகிறது. சங்கிலித் தொடர் (Chain base) அடிப்படை முறையில் குறியீட்டெண்கள் கணக்கிடப்படுகின்றன. (அதன்படி, 1975 என்ற ஆண்டினைக் கருதுகையில், அதன் அடிப்படை ஆண்டு 1974ஆக அமையும்).

(ii) விவசாய உற்பத்திபற்றிய, ரிசர்வ் வங்கியின் குறியீட்டெண்கள் (Reserve Bank of India Index) :

ஒவ்வொரு ஆண்டிற்கும், 17 பொருள்களை எடுத்துக்கொண்டு இக்குறியீட்டெண் கணக்கிடப்படுகிறது. ரிசர்வ் வங்கியின் மாதாந்திரச் சஞ்சிகையில் (Monthly) ஒவ்வொரு டிசம்பர் மாதத்திலும் இது வெளியிடப்படுகிறது. 1948-49 என்பது அடிப்படை ஆண்டாகக் கொள்ளப்படுகிறது. ஒவ்வொரு பொருளின் உற்பத்தி அளவும் எடையாகக் (Weight) கொள்ளப்பட்டு, எடையிட்ட சராசரி முறையைப் பயன்படுத்தி, குறியீட்டெண் கணக்கிடப்படுகின்றது.

(iii) ஐக்கிய நாடுகள் அமைப்பு (United Nations Organisation) மூலமாகவும், வெவ்வேறு நாடுகளின் விவசாய உற்பத்திக் குறியீட்டெண்கள் கணக்கிடப்பட்டுத் தொகுத்து வெளியிடப்படுகின்றன. இதன் மூலம், விவசாய உற்பத்தி நிலையைப்பற்றி வெவ்வேறு நாடுகளினிடையே ஒப்பீடு (Comparison) நடத்த முடிகிறது.

விவசாயத் துறையில் சில இதரப் புள்ளி விவரங்கள் :

(1) பயிர்கள் பற்றிய விவரங்கள் மட்டும் அல்லாமல், விவசாயத்தில் பயன்படுத்தப்பெறும் கால்நடைகள், கோழிகள் (போன்ற பறவைகள்) (Livestock and Poultry) பற்றிய விவரங்களும் சேகரிக்கப்பட்டு வெளியிடப்படுகின்றன. மாநில அரசினரால், ஐந்து ஆண்டுகட்கு ஒருமுறை (Quinquennial) இவற்றைப்பற்றிய முழுக் கணிப்பு நடத்தப்பட்டு, விவரங்கள் பெறப்படுகின்றன. மத்திய பொருளாதார மற்றும் புள்ளியியல் இயக்க அமைப்புகள் (Directorate of Economics & Statistics) இதற்கு உதவி புரிகின்றன. முதன்முதலில், 1920ஆம் ஆண்டு இவற்றைப்பற்றி கணிப்பு நடத்தப்பட்டது. இதுவரை சுமார் 10 கணிப்புகள் நடத்தப்பட்டுள்ளன. (1961ல் நடைபெற்ற கணிப் பின்படி, ஏறக்குறைய 4·5 கோடி அளவு கால்நடைகள் உள்ள தாக அறியப்படுகிறது.) தக்கமுறையில் வழிமுறைகள் திட்ட மிடப்பட்டு, கிராம, மற்றும் நகர்ப்புறங்களில் ஆய்வுகள் நடத்தப் பட்டு, தக்க மேற்பார்வையுடன் முடிவுகள் பெறப்படுகின்றன.

(2) அடுத்து, வனவளம் பற்றிய முழு விவரங்களும் பெறப் படுகின்றன. இதன்படி, வனப்பிரதேசத்தின் மொத்த நிலப் பரப்பு (Area of forest), கிடைக்கும் மொத்த மரத்தின் அளவு (Volume of timber), வனத்துறையினின்று கிடைக்கும் பொருள் களின் அளவு, பணியாற்றுவோர், வரவு செலவுநிலை, இப்பொருள் களின் வெளிநாட்டு வர்த்தக நிலை போன்ற பலவகையான புள்ளி விவரங்களும் சேகரிக்கப்படுகின்றன. இதற்காக, மத்திய அரசின் ஆண்டு வெளியீடான (Annual), 'இந்திய வனப் புள்ளி விவரங்கள்' (Indian Forest Statistics) என்பது வனவளத் துறை பற்றிய தேவையான விவரங்கள் அனைத்தையும் வெளி யிடுவதாக அமைகிறது. C. S. O. மூலமாகவும், (Statistical Abstract of India) இவை வெளியிடப்படுகின்றன. தகுந்த வழிமுறைகளில் விவரங்கள் திரட்டப்பட்டு, ஆய்வு நடத்தப்பட்டு, பயன்படத்தக்க முடிவுகள் பெறப்படுகின்றன.

(3) இறுதியில், மீன்வளத் துறை பற்றியும் சில விவரங்கள் அமைவதைக் குறிப்பிடுவது நன்று. இத்தகைய புள்ளிவிவரங்கள் முக்கியமாக, 'இந்தியக் கூட்டமைப்பின் மீன் வர்த்தக அறிக்கை, ('Report on Marketing of Fish in India') என்ற மத்திய அரசின் சந்தை மற்றும் ஆய்வு (Marketing and Inspecting) பற்றிய அலுவலகத்தின் மூலம் வெளியிடப்படுகின்றன. கிடைக்கும் மீன் களின் அளவு, பயன்படும் அளவு (Consumption), பயன்படும் விதம் போன்ற எல்லா விவரங்களும் வெளியிடப்படுகின்றன. கடலில் கிடைக்கும் மீன்கள் (Marine fish) பற்றிய மாதிரி ஆய்வு களும் விரிவான அளவில் மேற்கொள்ளப்படுகின்றன. 'இந்திய மீன்வளத் துறை வெளியீட்'டின் (Indian Journal of Fisheries) மூலமாக ஆய்வு விவரங்கள் வெளிவருகின்றன. NSS ஆய்வு களும் (வீடுகளில் பயன்படும் மீன்கள்பற்றிய அளவில்) பயன் படத்தக்க முடிவுகளை அளிக்கின்றன.

எனவே, மேற்கண்ட அனைத்தும் விவசாயத் துறையின் தொடர்புடைய சில பிரிவுகளாகக் கருதப்படுகின்றன.

IV. தொழிற் புள்ளி விவரங்கள் (Industrial Statistics)

ஒரு நாட்டின் பொருளாதார வளர்ச்சியானது, விவசாயத் துக்கு அடுத்தபடியாக, தொழில் வளர்ச்சியைப் பொறுத்தே அமை கின்றது. தொழில் மயமாக்கும் நடவடிக்கைகள் (Industrialisation) தற்போது பல்வேறு நாடுகளிலும் மேற்கொள்ளப்படு கின்றன. பொருளாதார வளர்ச்சியின் குறியீட்டு எண்ணாக (index) தொழில் நிலை விளங்கிடுகிறது. எனவே, தொழில் நிறுவனங்களின் அளவு நிலை (Size), முதலீடு (Capital) பணியாளர், கொடுக்கும் மொத்த ஊதிய அளவு (Total wages), உற்பத்தி, மூலப் பொருட்கள் போன்ற பல்வேறு இனங்களின் முழு விவரங்களையும் அறியாத நிலையில், தொழில் வளர்ச்சிக்குத் தக்க நடவடிக்கை எடுக்க இயலாது.

எனவே (i) வேலைவாய்ப்பு புள்ளிவிவரம்; (ii) முதலீட்டு நிலைபற்றிய புள்ளி விவரம் ; (iii) முதற்பொருள் (Input) பற்றிப் புள்ளிவிவரம் ; (iv) தொழில் உற்பத்தி அல்லது ஆக்கம் (Output) பற்றிய புள்ளிவிவரம் ; (v) கை இருப்பு (Stock), இலாப அளவு, சக்தி விரிவுபடுத்துதல் போன்ற பல்வேறு இதர விவரங்களைப் போன்ற பலவும் இப்பிரிவில் இடம்பெறுகின்றன. மேலும், சிறு தொழில்கள், குடிசைத் தொழில்கள்பற்றிய விவரங் களும் இதில் அடங்குவது குறிப்பிடத்தக்கது.

துவக்க நிலை (Origin)

தொழில் துறைபற்றிய புள்ளிவிவரங்களுக்கு, 'தொழில் புள்ளி விவரச் சட்டம்-1942' (Industrial Statitics Act-1942) அமைக்கப்பட்ட பின்பு அதிக முக்கியத்துவம் நம் நாட்டில் அளிக்கப்பட்டு வருகிறது. இதன்மூலம், வெவ்வேறு பிராந்திய அரசுகள், தொழில் உற்பத்தி நிலைபற்றி ஆய்வுகளை நடத்தி, தொழில்பற்றிய பல விவரங்களைப் பெற உரிமைபெற்றன. உற்பத்தித் தொழில்கள் (Manufacturing Industries) பற்றி முதன்முதல் 1946-ல் ஆய்வு நடத்தப்பட்டது. 1951முதல், தொழில்கள்பற்றிய விரிவாய்வுகளை, மாதிரி அடிப்படையில், NSS மேற்கொள்ளத் துவங்கிற்று. உற்பத்தியில் ஈடுபடும் எந்தக் குழுமத்தினின்றும் (Company) தேவையான விவரங்கள் பெறவேண்டியதன் அவசியத்தை உணர்ந்த மத்திய அரசு, 1953-ல், 'புள்ளிவிவர சேகரிப்பு சட்ட'த்தை(Collection of Statistics Act) உருவாக்கிற்று. இதன் பின்னர், உற்பத்தியில் ஈடுபடும் தொழில்களின் முழு ஆய்வும், மாதிரி ஆய்வும் ஒன்றுக இணைந்து, 'தொழில்கள்பற்றிய ஆண்டுக்கால ஆய்வு' என்ற விதமாக அமைக்கப்பட்டது.

தொழில் புள்ளிவிவரங்கள் சேகரிக்கப்பட, பின்வரும் சில ஆய்வு அமைப்புகள் உதவுகின்றன. அவையாவன :—

(I) உற்பத்தியில் ஈடுபடும் ஆக்கத் தொழில்களின் முழு ஆய்வு : (Census of Manufacturing Industries)

1946-ல் துவங்கி, ஆண்டுதோறும் நடைபெறும் ஒன்றுகும் இது. கட்டாய அடிப்படையில் (Statutory), 1956 வரையில் (11 ஆய்வுகள்) இது நடத்தப்பட்டது. முடிவுகளும் வெளியிடப்பட்டன. ஆய்வுக்கு, மொத்தம் பிரிக்கப்பட்ட தொழில்களுள் 29 மட்டுமே எடுத்துக் கொள்ளப்பட்டன. பொது விவரங்கள், முதலீட்டுத் திட்டம் (Capital structure), மொத்தப் பணியாளர் மற்றும் அவர்தம் ஊதியநிலை, எரிபொருள் உபயோகம், உற்பத்திப்பொருள்கள் மற்றும் துணைப் பொருள்களின் அளவு, போன்ற புள்ளிவிவரங்கள் பெறப்பட்டன. ஆனல் இம்முறையில், பெரும் அளவில் பதிலின்மை அமைந்தும், போதிய அளவு இடப்பரப்புக் கருதப்படாததாகவும், முடிவுகள் வெளியிடாமதம் ஏற்படுவதாகவும் அமைந்ததால், இந்த முறை ஆய்வானது கைவிடப்பட்டது.

(II) உற்பத்தித் தொழில்களின் மாதிரி அளவெடுப்பு (Sample Survey of Manufacturing Industries)

1951 முதல், இத்தகைய ஆய்வானது, NSS செயலகத்தின் (Directorate) மூலம் நடத்தப்பட்டது. தொழிற்சாலைகள் சட்டம்

(Factories Act—1948) மூலம் பதிவு செய்யப்பட்ட எல்லா நிறு வனங்களிலும் ஆய்வு நடைபெற்றது. பின்பு, தொழில்கள் ஒழுங்குமுறைச் சட்டம் (1951) மூலம் பதிவுபெற்ற நிறுவனங் களும் ஆய்வில் இடம் பெற்றன. முதலீடு, வேலை செய்வோர் மற்றும் ஊதியம், முதற்பொருள், உற்பத்தி போன்ற விவரங்கள் பெறப்பட்டன. 1958 முதல் இத்தகைய ஆய்வு முறை மேற் கொள்ளப்பட்டது.

(iii) தொழில்கள் பற்றிய வருடாந்திர ஆய்வு (Annual Survey of Industries)

1959-ல் ஏற்படுத்தப்பட்டு, முன் இரு வழிகளிலும் ஒரே விவரம் திரும்பவும் இடம் பெறுவதாக அமையும் நிலையினை நீக்க வழி ஏற்படுத்தியது. பாதுகாப்புத் துறையில் இடம்பெறும் நிறுவனங்கள், எண்ணெய் சேகரிப்பு நிலையங்கள், தொழிற் பயிற்சிக் கூடங்கள் இவற்றைத் தவிர ஏனைய தொழிற்கூடங் களில் ஆய்வு நடத்தப்பட்டது. மின் சக்தி (அல்லது திறன்) யுடன், 50 க்கு மேற்பட்ட பணியாளரைக் கொண்டு நடைபெறும் தொழில்களில் முழு ஆய்வும், அதற்குக் கீழ் உள்ள நிலையில் மாதிரி ஆய்வும் மேற்கொள்ளப்பட்டு, ஒவ்வொரு ஆண்டிலும் முதலீட்டுத் திட்டம், வேலை செய்வோர் மற்றும் ஊதிய அளவு, முதற்பொருள், உற்பத்தி, மின்சக்தி (திறன்) சாதனங்கள் (power equipments) போன்ற பலவற்றைப்பற்றி விவரங்கள் திரட்டப்பட்டன. ஒவ்வொரு தொழிலுக்கும், மாநிலத்துக்கும் தனித்தனியாக விவரங்கள் பெறப்பட்டன. மேலும், திறமை யுடைய (skilled) பணியாளர், திறமையற்ற பணியாளர், மின் ஆக்க சாதனங்கள் தவிர இதர சாதனங்களின் நிலை, உற்பத்தித் திறன், தொழில் ஆராய்ச்சி போன்ற பல கருத்துக்களுக்கு மிக முக்கியத்துவம் அளிக்கப்பட்டு, விரிவான அளவில் விவரங்கள் திரட்டப்பட்டன.

(iv) தேர்ந்தெடுக்கப்பட்ட சில தொழில்களில் மாதாந்தர உற்பத்திப் புள்ளி விவரங்கள் (Monthly Statistics of Production of Selected Industries)

இவை C.S.O. (மத்தியப் புள்ளிவிவரக் கழகம்)வின் தொழில் துறைப் பிரிவின் மூலம் மாதந்தோறும் வெளியிடப்படுகின்றன. இந்தத் திட்டத்தில்(scheme)இடம்பெறும் உற்பத்தி நிறுவனங்கள் தாமாகவே முன்வந்து (voluntarily) தரும் விவரங்கள் வெளி யிடப்படுகின்றன. இதில், சர்க்கரை, பருத்தி, சணல் ஆலைகள், இரும்பு மற்றும் எஃகு, தாதுப் பொருள்கள் போன்றவற்றைப் பொறுத்த அளவில், விவரங்களைப் பெறுவதற்கு, ஆய்வு

அதிகாரிகளுக்கு அதிகாரம் தரப்படுகிறது. முக்கியத் தொழில்களைப்பற்றிய முழு விவரங்கள், தேசீய அளவில் பெறப்படுகின்றன தேசீயத் தொழில் மற்றும் பணிகள் (occupational) அடிப்படையில் அமைந்த பாகுபாட்டிற்கு (classification) ஏற்ற விதத்தில் வெவ்வேறு தொழில்கள் பிரிக்கப்பட்டு, ஆய்வு நடத்தப்பட்டது. திறன் அளவு (capacity), பெரும்பாலான தொழில்களுக்கு விவரமாகப் பெறப்பட்டது. மாநில அளவில் மாதாந்திர உற்பத்தி அளவுகளும் சேர்க்கப்பட்டன. மேலும், 1960-ஐ அடிப்படையாகக் (base) கொண்டு, தொழில் உற்பத்திக் குறியீடும் (index) வெளியிடப்படுகிறது.

(v) மாதாந்திரப் புள்ளி விவரத் தொகுப்பு (Monthly Abstract of Statistics)

C.S.O.வின் வெளியீடான இதில் வெளிவரும் முக்கிய விவரங்கள்: (1) உணவு உற்பத்தி, துணி உற்பத்தி (textiles) இரசாயனப் பொருள்கள்; மரப் பொருள் உற்பத்தி: காகிதப் பொருள் உற்பத்தி; ரப்பர் பொருள்கள்; பெட்ரோலிய பொருள்கள் போன்ற இனங்களின் உற்பத்தி விவரங்கள்;

(2) எரிபொருள் மற்றும் மின் ஆற்றல் (Feul & Power) துறைகளில் உற்பத்தி மற்றும் பங்கீடு பற்றிய விவரங்கள்;

(3) மூலப் பொருள்கள் உபயோகிக்கப்படும் விவரம்; கையிருப்பு நிலை; உற்பத்தி செய்யப்பட்ட பொருள்களின் நிலை;

(4) தொழில் உற்பத்திக் குறியீடு போன்றவை ஆகும்.

(vi) இந்திய நாட்டின் புள்ளிவிவரத் தொகுப்பு (Indian Abstract of Indian Union)

இது CSOவின் வருடாந்திர (annual) வெளியீடாக அமைகிறது. தொழில் மற்றும் ஆற்றல் (Industry and Power) பற்றிய பல்வேறு விவரங்களும் வெளியிடப்படுகின்றன. இந்திய ரிசர்வ் வங்கியின் (Reserve Bank of India) மூலம் வெளியிடப்படும் 'பணம் மற்றும் நிதிநிலை பற்றிய அறிக்கை' Report on Currency & Finance) மற்றும் 'தொழில் வர்த்தக வெளியீடு' (Journal of Industry & Trade—வணிக அமைச்சகத்தின் மூலம் வெளியிடப்படுவது) ஆகியவற்றுடன் கால் வருடப் புள்ளி விவரங்கள்' (Quarterly Statistics) (முதலீடுபற்றிய விவகாரங்களைக் கவனிக்கும் பிரிவு வெளியிடுவது) போன்ற வெளியீடுகள் மூலமாகத் தொழில்துறைபற்றிய விவரங்கள் பெறப்படும். துணி உற்பத்தி (அல்லது சவுளி) பற்றிய அலுவலகம் வெளியிடும்

மாதாந்திர உற்பத்திப் புள்ளி விவரங்கள், (Monthly Statistics of Production) மூலமாக, பருத்தி உற்பத்தி போன்ற விவரங்கள் சேகரிக்கப்படுகின்றன.

சிறு தொழில், மற்றும் குடிசைத் தொழில் பற்றிய புள்ளி விவரங்கள் போதிய அளவு பெறப்படுவதில்லை. பெரும் அளவில் அமைந்த நிறுவனங்கள் 10 சதவீதமும், சிறுதொழில்கள் 8 சத வீதமும், தேசீய வருமானத்திற்கு ஆக்கப் பொருள்களை அளிப்பனவாய் அமையும் நிலை காணத்தக்கது. எனவே, சிறுதொழில்களைப் பொறுத்த அளவில், அவற்றின் திறன் அளவு, முதலீடு, வேலைவாய்ப்பு நிலை, ஆக்கம் போன்ற பல்வேறு விவரங்களும் முழு அளவில் பெறப்படுவது அவசியமாகிறது. மக்கட் கணிப்பு ஆய்வுகளில் அமைக்கப்படும் அட்டவணைகள் மூலம், இத் துறையில், ஓரளவு விவரங்கள் கிடைக்கின்றன. எனினும், மக்கட் கணிப்பு ஆய்வைப் பொறுத்த அளவிலும், NSS-ன் பல சுற்று ஆய்வுகளைக் கருதிய அளவிலும், போதுமான அளவு விவரங்கள் திரட்டப்படவில்லை என்றே கூற வேண்டும். சில ஆய்வுகளின் முடிவுகளை, திட்டக் கமிஷனின் (Planning Commission) ஆராய்ச்சி செயல்திட்டக் குழுவும் (Research Programme Committee), உணவு விவகார அமைச்சகமும், அவ்வப்போது வெளியிடுகின்றன.

சில துறைகளிலிருந்தும் (sources) விவரங்கள் தொடர்ந்து பெறப்படுவதைக் கூறலாம்.

அவையாவன:—

(i) தேசீய சிறு தொழிற்கள் கூட்டமைப்பின் (National Small Industries Corporation) செயல்முறைபற்றிய வருடாந்திர அறிக்கைகள்;

(ii) வெவ்வேறு மாநிலங்களிலும் இத்தகைய கூட்டமைப்புகளின் பணிமுறை பற்றிய வருடாந்திர அறிக்கைகள்;

(iii) ரிசர்வ் வங்கி வெளியீடான பணம் மற்றும் பொருளாதார நிலைபற்றிய ஆண்டறிக்கை. இதில், நிதி உதவி பெறும் பல்வேறு சிறு தொழில் நிறுவனங்களைப்பற்றிய விவரங்களும், ஸ்டேட் (மாநில) வங்கிகள் (State Bank) உதவும் நிதி அளவு பற்றிய விவரங்களும் அடங்குகின்றன.

தொழில் உற்பத்திக் குறியீடு (Index of Industrial Production)

தொழிற் புள்ளி விவரங்கள்பற்றிய இயக்க அமைப்பின் (directorate) மூலம், உற்பத்திப் பொருள்கள் 324ஐக் கொண்டு

மதிப்பிடப்படும் குறியீட்டு எண் ஒவ்வொரு மாதமும் வெளியிடப் படுகிறது. 1960 ஆண்டானது (= 100) அடிப்படையாக (base) கொள்ளப்பட்டு இக்குறியீடானது கணக்கிடப்படுகிறது. இதோடு கூட, ஒவ்வொரு ஆண்டிலும், 125 துணைப் பொருள்களைப்பற்றிய விவரங்களும் வெளிவருகின்றன. இந்தியத் தொழில்களின் தரப் பாகுபாட்டின் (Indian Standard Industrial Classification) அடிப்படையில், இக்குறியீட்டில் இடம்பெறும் பொருட்கள் (items) பிரிக்கப்படுகின்றன. உற்பத்தி மூலம் அதிகரித்த மதிப்புக் களை(values) எடைகளாகக் கொண்டு, உற்பத்தித்தொடர்பீடு களின் (Production relative) எடையிடப்பட்ட சராசரி முறை மூலம் குறியீடானது கணக்கிடப்படுகிறது. பருவ மாறுபாட்டுக்கு ஏற்ப வெவ்வேறு மாதங்களின் குறியீடுகள் திருத்தப்பட்டு வெளியிடப்படுகின்றன.

பொதுவாக, தொழிற் புள்ளிவிவரங்களைப் பொறுத்த அளவில் நம்நாட்டுச் சூழ்நிலையில், பின்வரும் குறைபாடுகளைக் காணலாம் :

(1) ஒரு சில முக்கியமான தொழில்களைக் கருத்திற்கொண்டு புள்ளி விவரங்கள் பெறப்படுவதால், அவை விரிவான (exhaustive) அளவில் அமைந்திருப்பதில்லை. வரிகளை ஏய்க்கும் பொருட்டு (tax evasion), தொழில் நிறுவனங்கள் தமது உண்மை யான நிலைபற்றிய விவரங்களை அளிப்பதில்லை. விவரங்கள் தவறாக இருப்பதால், முடிவுகளும் நம்பத்தக்கனவாக இருப்ப தில்லை; தேவையான விவரங்களைப் பெறுவதற்கு தகுந்த விதி முறைகளும் அமைக்கப்படவில்லை.

(2) விவரங்களை வெளியிடுவதில் மிகவும் கால தாமதம் ஏற்படுவதால், தற்காலச் சூழ்நிலையைப் பிரதிபலிக்கும் அளவில் முடிவுகள் பெற இயலுவதில்லை. மேலும், ஆய்வுகளில் பயன் படுத்தப்படும் வெவ்வேறு சொற்களும் கருத்துக்களும் வரையறை களும் தெளிவாக விளக்கப்படாததால், அவை அனைத்தையும் மாற்றி எளிதாக அமைக்கவேண்டியது இன்றியமையாததாகும். பொதுவாக, விலையைக் குறிப்பிடும்போது, 'சந்தை விலை' (market-price) குறிப்பிடப்படுவதா அல்லது 'உற்பத்தி விலை' (factory-price) கொள்ளப்படுவதா என்றவாறான ஐயம் ஏற்பட வழியுண்டு.

(3) தொழில் உற்பத்தியில் ஒரு பிரதான அல்லது முக்கிய இடம்பெறும் சிறு தொழில்கள் தகுந்த முறையில் ஆய்வுக்குக் கருதப்படவில்லை. எனவே, சிறு அளவில் உற்பத்திகள் மேற்

கொள்ளப்படும் தொழில் நிறுவனங்கள், குடிசைத் தொழில்கள் போன்றவற்றில் ஆய்வுகள் நடத்தவும், விவரங்கள் சேகரிக்கவும், விவரங்களைத் தக்க முறையில் தொகுத்து முடிவுகளை வெளியிடுவதற்கும் தகுந்த அளவில் முயற்சி எடுக்கப்படவேண்டியது மிக இன்றியமையாததாகும்.

எனவே, தொழிற் புள்ளி விவரங்களை முறையாகவும், போதுமான அளவிலும் பெறும் பொருட்டு, ஆண்டுதோறும் மேற்கொள்ளப்படும் தொழில்கள் பற்றிய விரிவாய்வானது (surveys of industries), தகுந்த முறையில் அமைப்புகள் அற்ற பிரிவில் அமையும் நிறுவனங்களையும் கருத்தில் கொண்டு செயல் படவேண்டியது அவசியமாகிறது. அந்நிலையில் நாட்டுப் பொருளாதார நிலை பற்றிய ஒரு தெளிவான முடிவினைப் பெறுவதற்கு நம்பத்தக்க முறையில் ஒரு வழி ஏற்படுகின்றது. இந்தியப் புள்ளி விவரங்கள் பற்றிய விரிவுப் பகுப்பாய்வானது, விவசாயம் மற்றும் தொழில் துறைகள் பற்றி இதுவரை தரப்பட்ட விவரங்களுடன், இனி எடுத்துக்கொள்ளும் 'வர்த்தகத்' துறை (trade) பற்றிய சில குறிப்புகளையும் உள்ளடக்கியதாக அமைவதாகும். அஃதாவது :—

V. வர்த்தகப் புள்ளி விவரங்கள் (Trade Statistics)

இந்திய வர்த்தகப் புள்ளி விவரங்களானவை இருபெரும் பிரிவுகளை உடையதாகக் காணப்படுகின்றன. அவை:

(1). உள் நாட்டு வர்த்தகம் (Internal Trade)

(2). வெளி நாட்டு வர்த்தகம் (Foreign Trade)

(1) முதல் வகையில், ஒரு நாட்டினைப் பொறுத்த அளவில், வெவ்வேறு மாநிலங்களின் (அல்லது வட்டாரப் பிரிவுகளின்-regions) இடையிலும், மற்றும் ஒவ்வொரு மாநிலத்தின் (அல்லது வட்டாரப் பிரிவின்) உள்ளும் (within) நடைபெறும், பொருட்களின் பரிவர்த்தனை (exchange of goods) மொத்த அளவினைக் குறிப்பதாக, 'உள்நாட்டு' வர்த்த நிலையானது அமைந்திருக்கிறது. இந்திய நாட்டின் அபரிமிதமான பொருளாதார வளர்ச்சியின்மூலமாக, உள்நாட்டு வர்த்தகமானது பெரும் அளவில் விரிவடைந்திருக்கக் காணலாம். எனவே, வெவ்வேறு வட்டாரப்பகுதிகளும் ஒன்றைஒன்று சார்ந்து விளங்கும் (inter dependence) தன்மையினையும், வெவ்வேறு மாநிலங்களினிடையில் நடைபெறும் வர்த்தக நிலையினையும் நன்கு

கண்டறிந்து முடிவுகளைப் பெறுவதற்கு, உள்நாட்டு வர்த்தகம் பற்றி முழு அளவிலும், நம்பத்தக்க விதத்திலும் உள்ள புள்ளி விவரங்கள் பெறப்படுவது மிகவும் முக்கியமானதாகும். இந்தியப் பொருளாதார நிலை பற்றிய மதிப்பீட்டினைப் பெறவும் இவை உதவுகின்றன எனலாம். ஒரு பகுதியிலிருந்து மற்றொரு பகுதிக்குப் பரிவர்த்தனை செய்யப்படும் பொருள்கள் பற்றி விரிவான அளவில் ஆய்வுகள் மேற்கொள்ளப்படுகின்றன.

குறிப்பாக, இரு விதங்களில் உள் நாட்டு வர்த்தகம் பற்றிய விவரங்களை

(i) 'மொத்த' வர்த்தக விவரங்கள் (Wholesale)

(ii) 'சில்லறை' வர்த்தக விவரங்கள் (Retail)

என்றவாறுகக் கொள்வது இயல்பு.

(i) 'மொத்த' வர்த்தகமாவது :

(அ) இருப்புப்பாதை மற்றும் நதி வழி வர்த்தகம் (Rail and River -borne Trade)

இருப்புப் பாதை வழியாக. (ரயில்வே அமைப்புகள் மூலமாக) மற்றும் பல்வேறு நதிகளின் வழியாக மேற் கொள்ளப்படும் உள் நாட்டு வர்த்தகங்கள் பற்றிய விரிவான புள்ளி விவரங்கள், கல்கத்தாவில் அமைக்கப்பட்டுள்ள வணிகத்துறைப் புலனாய்வு மற்றும் இயக்கச் செயலகத்தின் பற்றிய புள்ளி விவரம் Direcortate General of Commercial Intelligence and Statistics of India D. G. C. I. S) மூலமாக ஒவ்வொரு மாதமும் தொகுக்கப்பட்டு வெளியிடப்படுகின்றன இவ் விவரங்கள், நாட்டில் உள்ள எட்டுப் பிராந்திய ரயில்வேகள் (zonal railways) மூலமும், கப்பல்கள் மூலமும் (steamers) ஒவ்வொரு மாதமும் அளிக்கப்படுகின்றன. 30 முக்கிய வகைகளாக, வர்த்தகப் பொருட்கள் அனைத்தும் பிரிக்கப்படுகின்றன

இவற்றில் சில குறைபாடுகள் : முக்கியமாக, போதுமான அளவு விவரங்கள் பெறப்படுவதில்லை. ஒருசில கப்பல் கம்பெனிகளைக் (Steamer Companies) (கூட்டு நிறுவனங்களை) கருதுவதன்மூலம், நதி வழி வர்த்தக நிலைபற்றிய விவரங்கள் பெறப்படுவதால் நம்பத்தக்க விவரங்கள் அமையாத நிலை ஏற்படுகிறது. நாட்டில் பல்வேறு பகுதிகளின் படகுகளின் மூலம் நடைபெறும் வியாபாரமும் கணக்கெடுக்கப்படுவதில்லை. பொருட்களின் போக்கு வரத்து மட்டும் கணக்கில் கொள்ளப்பட்டு பார்சல்களின் போக்கு வரத்துக் கருதப்படுவதில்லை, எனவே தவறான மதிப்பீடுகள் பெறப்படுகின்றன. மேலும், வர்த்தகத்தில்

முக்கிய அம்சமான மதிப்பீடுகள் (valuation) காணப்படுவதில்லை, எனவே, விவரங்களை ஒப்பிட இயலாது போகிறது.

(ஆ) கடல்வழி வர்த்தகம் (Coastal Trade)

கடற்கரைப் பிரதேசங்களில் (உள்நாட்டு வாணிபம்) நடைபெறும் வர்த்தக நிலைபற்றிய புள்ளிவிவரங்கள், கல்கத்தாவிலுள்ள D.G.C.I. & S. (ஏற்கெனவே விளக்கப்பட்டது) மூலமாக, சேகரிக்கப்பட்டு, இதற்கெனஅமைந்த 'கடற்கரை வர்த்தக நிலைப் புள்ளி விவரங்கள் (இந்திய நாட்டில்)' (Statistics of the Coasting Trade in India) என்னும் வெளியீட்டின்மூலம் ஒழுங்கான காலவரையில் வெளியிடப்படுகின்றன. கடல் மார்க்கப் போக்குவரத்து (Navigation) வர்த்தகம் பற்றிய, மற்ரொரு வெளியீடான ('Statistics of Maritime Navigation of India') என்பதன் மூலமும் விவரங்கள் வெளியாகின்றன. அங்கங்கு கடற்கரைப் பிரதேசங்களில் உள்ள சுங்க அதிகாரிகள் (Customs Official) (இந்தியாவில்), வர்த்தக நிலை பற்றிய விவரங்களை அளிக்கின்றனர். இந்தியக் கடற்கரைப் பிரதேசங்களை, ஆய்வுக்கு எளிதாக, 12பகுதிகாகப் பிரித்து (Maritime Blocks) விவரங்களைக் கண்டறிகின்றனர்.

(ii) சில்லறை வர்த்தக (அல்லது வாணிப) நிலை

'இந்திய வர்த்தக வெளியீடு' (Indian Trade Journal), வணிகப் புலனாய்வு மற்றும்புள்ளிவிவரம் பற்றிய இலாகா மூலம் (Department of Commercial Intelligence & Statistics) ஒவ்வொரு வாரமும் வெளியிடப்பட்டு, நாட்டில் பல பகுதிகளிலும் நடைபெறும் சில்லறை வாணிப நிலைகளை (Retail Trade) விவரங்களுடன் அறிய எளிதாகிறது. ஆனால், இதுபோன்ற விவரங்கள் அனைத்தும் போதிய அளவிலும் நம்பத்தக்க அளவிலும் அமைந்திராதது பெரியகுறைபாடாகிறது. எனவே, பகுப்பாய்வுகட்கு இம் முடிவுகள் உதவி செய்வதில்லை.

எனவே, பொதுவாக உள்நாட்டு வர்த்தக (வாணிப) நிலை பற்றி விளக்கும் புள்ளி விவரங்களில் காணப்படும் குறைபாடுகள் பின்வருமாறு அமைகின்றன:

உள்நாட்டு வர்த்தகத்தில் ஈடுபடுத்தப்படும்பல்வேறு பொருட்களின் உண்மையான மதிப்பீட்டு எண்கள் (Value Figures) கிடைப்பதில்லை. அடுத்து, ஒவ்வொரு பொருளாக எடுத்துக் கொண்டு ஆய்வு நடத்தும் வழிமுறை பின்பற்றப்படுவதில்லை. சில்லறை வர்த்தகச்சூழ்நிலைபற்றி அதிக அளவில் விவரங்கள்

திரட்டப்படவில்லை. மேலும், அதிமுக்கியமான. தரைவழியாக கொண்டுசெல்லப்பட்டு வர்த்தகத்தில் ஈடுபடுத்தப்படும் பொருட்களைப்பற்றி எந்தவித விவரமும் பெறப்படுவதில்லை, உதாரணமாக, மாட்டுவண்டிகள்மூலம், அல்லது லாரி போன்ற (Trucks) உந்துக்கள்மூலம் அல்லது வேறெந்த வழியிலும் கொண்டு செல்லப்படும் சரக்குகள் (வாணிப நிமித்தம்), இந்தப் புள்ளி விவரங்களில் இடம் பெறுவதில்லை. இது முக்கியமாகக் கவனித்துக் களையத்தக்க குறைபாடாகும். மற்றும், கடற்கரை ஓரப் பகுதிகளில் நடத்தப்படும் ஆய்வுகளில், ஏற்றுமதிகளும் இறக்குமதிகளும் நடைபெறும் மூல அமைப்புகளும் (Sources), முடிவு நிலைப் பகுதிகளும் (Destinations) வர்த்தகத் தொகுதி (Trade block)களாகக் கருதப்படுகின்றன. எனவே, அடிப்படை விவரங்கள் சேகரிக்கப்படாவிடின், மாநில அளவில் மற்றும் வட்டார அளவில் நடைபெறும் வர்த்தகத் தன்மையை நன்கறிய இயலுவதில்லை. விண்வழிப் (Air) போக்குவரத்து (சரக்குகளைப் பொறுத்தவரை) பற்றிய விவரங்கள் திரட்டப்படுவதில்லை. எனவே பல்வேறு வழிகளிலும், இந்திய உள்நாட்டு வர்த்தக நிலைபற்றிய முழுவிவரங்கள் பெறப்படுவதில்லை என்பது தெளிவாகிறது.

ஓரளவு, நம் நாட்டில் நடைபெறும் (உள்நாட்டு) வர்த்தக நிலையினைப்பற்றி ஆராய்ந்தபின், வெளிநாட்டு வர்த்தகம் (Foreign trade) பற்றிய புள்ளிவிவரங்கள் எந்த நிலையில் உள்ளன என்று காணவேண்டிய அவசியம் ஏற்படுகிறது. உலக நாடுகளுடன் பெரும் அளவு பொருளாதார மற்றும் வாணிபத் தொடர்புகொண்டு விளங்கும் நம் நாட்டில், வெளிநாட்டு வர்த்தகமானது ஒரு சிறந்த இடத்தினைப் பெறுகிறது என்பதில் ஐயமில்லை. ஏற்றுமதியைப் பெருக்கி, இறக்குமதியைச் சுருக்கி, பொருளாதார வளத்தைச் செம்மைப்படுத்துதற்கு வழிமுறைகள் மேற்கொள்ளப்படுகின்றன.

(2) வெளிநாட்டு வர்த்தகம்பற்றிய புள்ளி விவரம் (Foreign Trade)

நாட்டின் சுங்க இலாகாவினரால் (Customs) கட்டுப்படுத்தப்பட்ட பகுதிகளில் இறக்குமதியாகும், மற்றும் அவற்றினின்று வெளிநாடுகட்கு ஏற்றுமதியாகும் பொருட்களைப் பற்றிய விவரங்கள் இத்தகைய புள்ளிவிவரங்களை உருவாக்குகின்றன. இத்தகைய விவரங்கள் பல்வேறு பயன்களைக் கொண்டன ஆகும். நாடுகளின் இடையில் பரிவர்த்தனைச் சமன்பாட்டை (Balance of Payments) நிர்ணயிப்பதோடு, பல்வேறு நாடு

களோடு நடத்தப்படும் வர்த்தகநிலைகள் பற்றிய தெளிவான விவரங்களைப் பெறவும், இறக்குமதியாகும் பொருட்களுக்கு உள் நாட்டில் எத்தகைய சந்தைநிலை அமைகிறது என்பதனை மதிப்பிட்டு அறியவும், வெளிநாட்டு வாணிபப் புள்ளி விவரங்கள் நன்கு உதவுகின்றன. ஏற்றுமதியை அதிகப்படுத்தவும், இறக்குமதிகளைக் கட்டுப்படுத்தவும் வழிமுறைகள் மேற்கொள்வதானது, இப் புள்ளி விவரங்களின் அடிப்படையில்தான் அமைகிறது. வெளிவர்த்தக நிலைபற்றிய ஒரு சிறந்த மதிப்பீட்டையும் பெற முடியும். நம் நாட்டைப்போன்ற (பொருளாதாரத் துறையில்) வளர்ந்துவரும் நாடுகளில் இது இன்றியமையாததாகும். வர்த்தக உறவுகளை (Trade agreements) ஏற்படுத்தவும், இப் புள்ளி விவரங்கள் வழி செய்கின்றன.

இந்திய நாட்டு வெளி வர்த்தக (foreign trade) புள்ளி விவரங்கள் அனைத்தும் (ஏற்கெனவே குறிப்பிடப்பட்ட) '*DGCI* and *S*' மூலமாகத் தொகுத்து வெளியிடப்படுகின்றன. சுங்க மற்றும் தீர்வை (Customs & Excise) சம்பந்தப்பட்ட பிரிவுகளின் மூலம் அளிக்கப்படும் விவரப் பட்டியலை (returns) கொண்டு, இப் புள்ளிவிவரங்கள் வெளியிடப்படும். வெளி வர்த்தக நிலையைப் பொறுத்தவரையில், புள்ளி விவரங்கள் போதிய அளவும், நம்பத்தக்கனவாகவும் அமைந்திருப்பது குறிப்பிடத்தக்கனவாகும். ஒரு சில வெளியீடுகளைக் குறிப்பிடுவோம் :

[அனைத்தும் அரசினரால் (official) வெளியிடப்படுபவை].

(i) 'இந்திய நாட்டின் வெளிநாட்டு வாணிப (அல்லது வர்த்தக) நிலை பற்றிய மாதாந்திரப் புள்ளி விவரம் :—

(Monthly Statistics of Foreign Trade)—Vol. I.

இம்முதற்பாகமானது (Vol. I) ஏற்றுமதிகள், மறு-ஏற்றுமதிகள் (Exports & Re-exports) பற்றிய விவரங்களை வெளியிடுவதாகும். நாட்டின் எல்லாத் துறைமுகங்களையும், விமான நிலையங்களையும், சுங்கச் சாவடிகளையும் கடந்து செல்லும் சரக்குகள் பற்றிய அனைத்து விவரங்களும் தரப்படுகின்றன. பொதுவாக, வெள்ளி, செலாவணியில்லாத நாணயங்களும் ரூபாய் நோட்டுகளும், பார்சல் மூலம் அனுப்பப்படும் ஏற்றுமதிப் பொருட்கள் போன்ற பலவும் இதில் இடம்பெறுகின்றன, இடம் விட்டு இடம் செல்லும் பொருள்கள் (வியாபார நிமித்தமாக இல்லாமல்), கப்பல்களின் உடைமைகள் போன்ற சில பொருட்கள் கணக்கில் சேர்க்கப்படுவதில்லை. 'வர்த்தகப்

பாகுபாட்டு விதி'க்கு ஏற்ப (Revised Indian Trade Classification) பொருட்கள் பல்வேறு வகையாகப் பிரிக்கப்படுகின்றன. வெவ்வேறு பிரிவுகளிலும், தொகுதிகளிலும், பகுதிகளிலும் (Groups) அமையும் ஏற்றுமதிகளைக் காட்டும் வண்ணம் விரிவுப் பட்டியல்கள் அமைக்கப்படுகின்றன. நாடுகளின் அளவிலும் ஏற்றுமதி மற்றும் மறு-ஏற்றுமதி பற்றிய விளக்கமான அட்டவணைகள் அமைக்கப்படுகின்றன.

(ii) மேற்கண்ட அதே வெளியீட்டின் இரண்டாம் பாகமானது (Vol.II), நாட்டின் இறக்குமதிப் பொருள்களைப் (Imports) பற்றியதாகும். எல்லா துறைமுகங்களுக்கும், விமான நிலையங்கட்கும், சுங்கச் சாவடிகளுக்கும் வந்து சேரும் பல்வேறு சரக்குகளும் (merchandise) கணக்கெடுக்கப்படுகின்றன. புழக்கத்தில் இல்லாத வெள்ளி, வெளியிடப்படாத நாணயங்கள், பார்சல் வழியாக வந்து சேரும் இறக்குமதிப் பொருட்கள், பாதுகாப்புத் துறைக்கு வாங்கப்படும் பொருட்கள் போன்றவை இடம் பெறுகின்றன. இதன் நிலையும், முதலில் கொடுக்கப்பட்டது போலவே அமைகிறது.

(iii) 'இந்திய வெளி வர்த்தகம் பற்றிய மாதாந்திரப் புள்ளி விவர' வெளியீட்டினது (Vol. I) பிற்சேர்க்கை (Supplement I) யாகச் சில துணை விவரங்கள் D.C.I.S. மூலம் ஒவ்வொரு காலாண்டிலும் (3 மாதத்திற்கு ஒருமுறை) வெளியிடப்படுவதாகும்.

இதில் பின்வரும் அட்டவணைகள் (பட்டில்கள்) வெளியாகின்றன:

(i) வெளிநாட்டு வாணிபத்தின் மதிப்பு (Value)

(ii) வர்த்தகத்தின் மொத்தமான சமநிலை (Balance)

(iii) சுங்கப் பகுதிகளின் (Customs zone) வெளி வர்த்தகம்.

(iv) இந்தியாவைக் கடக்கும் நாடுகளின் வெளி வர்த்தக நிலை.

(v) குறியீட்டெண்கள்.

(vi) முக்கிய ஏற்றுமதிப் பொருட்களின் மதிப்பும், அளவும்.

(vii) முக்கிய இறக்குமதிப் பண்டங்களின் மதிப்பும், அளவும்.

(viii) நிதியகத்தின் (Treasure) மொத்த வெளி வாணிபம்.

இவை முக்கியமானவையாகக் கொடுக்கப்பட்ட, துணை நிலை விவரங்களாகும்.

(iv) 'Monthly Statistics of Foreign Trade of India, Volume II (இரண்டாம் பாகம்) என்பதன் (பிற்சேர்க்கை வெளியீட்டில் (II), இந்திய நாட்டுடன் இதர உலக நாடுகள் மேற்கொள்ளும் (பண்டங்கள் பற்றிய) வர்த்தக நிலை விவரங்கள் வெளியாகின்றன. இப்புள்ளி விவரங்கள் அனைத்துமே அகில உலக வர்த்தகத் தரப் பாகுபாட்டை (Standard International Trade Classification) அடிப்படையாகக் கொண்டு வெவ்வேறு பகுதிகட்கும், இனங்கட்கும் வெளியாகின்றன. இவ்வெளியீட்டில் (Supplement II), சில விவரங்களாவன:

(a) நாடுகள் அளவில் மற்றும் பொருளாதாரப் பிரிவுகள் அடிப்படையில், இந்திய வெளி வர்த்தக நிலை.

(b) முக்கியப் பண்டங்களைப் பொறுத்து, பின்வரும் நாடுகளுடன் நடைபெறும் வாணிபத் தன்மைபற்றிய விவரங்கள் :

1. அமெரிக்க நாடுகள்;
2. ECAFE நாடுகள்
3. மற்ற ஆசிய நாடுகள்;
4. கிழக்கு ஐரோப்பிய நாடுகள்;
5. ஆப்பிரிக்கா
6. E.C.M-ல் இடம்பெறும் நாடுகள்.
7. இதர ஐரோப்பிய நாடுகள்.

இவ்வெளியீட்டின் மூலம், வெளிநாட்டு வர்த்தகச் சூழ்நிலையினைத் தெளிவுற அறிய முடிகின்றது.

(v) இந்திய வெளி வர்த்தகம் பற்றிய ஓராண்டுப் புள்ளி விவரம் (ஆண்டுதோறும்) (Annual Statistics of Foreign Trade of India) :

மாத வெளியீடுகளில் வெளியிடப்படும் வெளி வர்த்தகப் புள்ளி விவரங்களின், மாதாந்திரக் கூட்டு மொத்த மதிப்புகளை (monthly totals) கொண்டு அமையும், சரிபார்க்கப்பட்ட, மறு பரிசிலனை செய்யப்பட்ட ஒன்றுகும் இவ் வெளியீடு. இதில் பின்வரும் விவரங்கள் இடம்பெறும் :

- ☐ ஒவ்வொரு நாட்டுடனும் நடைபெறும் மொத்த வர்த்தக அளவு;
- ☐ ஒவ்வொரு நாட்டிற்கும் தனித்தனியாகக் கணக்கிடப்பட்ட, பொருள்வாரி விவரங்கள் ;
- ☐ எரிபொருள் (fuel) ஏற்றுமதி அளவு ;

☐ வெளிநாட்டு மற்றும் உள்நாட்டுப் பரிவர்த்தனைமூலம் நிகழும் வாணிப அளவு;

(vi) சுங்க மையங்களுக்கு (Customs zones) வெளியிடப்படும் ஆண்டு வெளியீடு (Annual Statistics of Foreign Trade by Customs Zones) :

கல்கத்தா, சென்னை, டெல்லி, கொச்சி, பம்பாய், பாட்னா, பரோடா, ஷில்லாங் போன்ற முக்கியமான மையங்களைப் பொறுத்த அளவில் நடைபெறும் இந்திய வெளி வர்த்தகம் பற்றிய விவரங்களை அளிப்பதாகும் இவ் வெளியீடு. இதனை, முன்பு குறிப்பிட்ட ஆண்டு வெளியீட்டோடு இணைப்பதற்குத் திட்டமிடப்பட்டுள்ளது.

(vii) இந்தியக் கூட்டமைப்பின் சுங்க மற்றும் தீர்வை வருவாய் பற்றிய அறிக்கைகள் (Customs & Excise Revenue Statements) :

இவை ஒவ்வொரு மாதமும் வெளியாகின்றன. பல்வேறு வகையான சுங்கவரி, எக்ஸைஸ் தீர்வை போன்ற வரிகள்மூலம் கிடைக்கும் வருவாய் விவரங்கள் அனைத்தும் தரப்படுகின்றன. சுங்க மற்றும் தீர்வை பிரிவுகள் தொடர்புடைய அதிகாரிகள்மூலம் இவை *DCIS*-க்கு அளிக்கப்பட்டு, பின்பு விவரங்கள் தொகுத்தளிக்கப்படுகின்றன.

(viii) இந்திய வெளி வர்த்தகத்தின் வருடாந்திரப் பிற்சேர்க்கை வெளியீடு: (Annual Supplement to Foreign Trade of India)

சிறு அளவில் கப்பல் மூலம் ஏற்றுமதியாகும் பண்டங்களின் அடிப்படையில், வெளி வர்த்தக மாதாந்திர அளவுகளைச் சரி பார்த்து, திருத்தங்களுடன் மதிப்புகள் வெளியிடப்படுகின்றன. இதேபோல் இறக்குமதி விவரங்களும் சரிபார்க்கப்படுகின்றன. ஒவ்வொரு நாட்டுடனும் நடத்தப்படும் வாணிப அளவு, விரிவான முறையில் வெளியிடப்படும்.

(ix) தொழில், வர்த்தக வெளியீடு (Journal of Industry & Trade)

மத்திய வாணிக அமைச்சகத்தின் விளம்பரப் பிரிவு (Commercial Publicity) மாதந்தோறும் வெளியிடுவதாகும். இதன் ' வெளி வர்த்தகப் பகுதி'யானது, 'முக்கியப் பொருட்களின் ஏற்றுமதி,

முக்கியப் பண்டங்களின் இறக்குமதி, மொத்த வாணிப அளவு, வர்த்தகத்தின் போக்கு ஆகிய விவரங்களைக் கொண்டதாக அமைகிறது.

(x) இந்தியப் புள்ளிவிவரத் தொகுப்பு (Indian Statistical Abstract)

CSO-வின் வெளியீடான இது ஆண்டுதோறும் வருவதாகும். கடற்கரை வாணிபம் (Coastal trade), கடல் வழி மற்றும் விண் வழி வர்த்தகம், உள்நாட்டு வாணிபம், போன்ற விவரங்கள் இதில் இடம் பெறுகின்றன.

மேற்குறிப்பிட்டவை முக்கியமான சில வெளியீடுகளாகும். Statistics of Maritime Navigation of India-Annual, The Inland Trade; Reserve Bank of India Monthly Bulletin போன்றவற்றையும் சேர்த்துக் கொள்ளலாம்.

ஏற்றுமதி இறக்குமதிக் குறியீடுகள் (Indexes of Exports and Imports)

இந்திய வெளி வர்த்தக நிலை பற்றிய குறியீட்டெண்கள் (Index numbers), D. C. I. S. மூலம் தொகுத்து அளிக்கப் படுகின்றன. தனித்தனியாக வெவ்வேறு பண்டங்களையும் பாகுபாடு செய்து, ஏற்றுமதி மற்றும் இறக்குமதி பற்றிய குறியீடுகள் கணக்கிடப்படுகின்றன. ஏற்றுமதி இறக்குமதி அளவுகள் குறித்தும், மொத்த அளவின் மதிப்பு குறித்தும் (volume) இத்தகைய எண்கள் கணக்கிடப்பட்டு வெளியிடப் படுகின்றன. இவை அனைத்தும், 'இந்திய வெளி வர்த்தக மாதாந்திரப் புள்ளி விவரத் துணை வெளியீட்'டில் (Supplement) இடம் பெறுகின்றன.

மொத்தம் 511 இறக்குமதிப் பொருள்களும், 317 ஏற்றுமதிப் பொருள்களும் இக் குறியீடுகள் கணக்கிடுவதற்கு எடுத்துக் கொள்ளப்பட்டன, மேலும், மறு ஏற்றுமதிப் பொருட்கள் (re-exports) தனியாகக் கருதப்பட்டு, கணக்கீடு எளிதில் அமைய உதவியது. தரமான இந்திய வர்த்தகப் பாகுபாட்டு முறையின்படி (Standard Indian Trade Classification — 1957) அடிப்படை ஆண்டானது 1958 என்று தேர்ந்தெடுக்கப்பட்டு, குறியீட் டெண்கள் கணக்கிடப்பட்டு வருகின்றன.

நிலையான அடிப்படை ஆண்டைப் (fixed base) பயன் படுத்தி, வர்த்தகப் பொருட்களின் அளவுபற்றிய குறியீட்

டெண்கள் பின்வருமாறு பெறப்பட்டன; P என்ற வர்த்தகக் குறியீடானது,

$$P = \frac{\Sigma p_1 \cdot q_1}{\Sigma p_0 q_0}$$

ஆகும். இதில் p_0, p_1, ஆகியவை அடிப்படை ஆண்டிலும் நிகழும் ஆண்டிலும் உள்ள (பொருள்களின்) விலைகளாகவும், q_0, q_1 என்பவை இக்காலங்களுக்கு முறையே அமைந்த அளவுகளாகவும் (quantities) வரையறுக்கப்படுகின்றன.

மொத்த வர்த்தகக் குறியீடு (quantum)

$$Q = \frac{1}{P} \left(\frac{V_1}{V_0}\right) \times 100 \text{ ஆகும்.}$$

V_0, V_1 என்பவை வர்த்தக மொத்த மதிப்புகள் (Values) என்றாகின்றன.

இந்திய நாட்டுப் புள்ளி விவரங்கள்—ஒரு திறனாய்வு.

நம் நாட்டில் சேகரிக்கப்பட்டு ஆய்வு மேற்கொள்ளப்படும் சில முக்கியமான, அடிப்படைப் புள்ளி விவரங்கள் பற்றிய சில பொதுவான இயல்புகளை இங்கு விவரிப்போம் :

பெறப்படும் புள்ளி விவரங்கள் போதிய அளவில் இருப்பதில்லை. அதாவது, எந்த ஒரு ஆய்வுக்கும் பயன்படும்படியாக, முழு அளவில் விவரங்கள் திரட்டப்படுவதில்லை. மத்தியப் புள்ளி விவர ஆய்வுக் கழகமானது (*CSO*) தகுந்த நடவடிக்கைகள் எடுத்து, இக் குறையினைப் பெரும் அளவு களைய முற்படுகிறது. பொருளாதாரத்தின் பல பிரிவுகள் பற்றிய முழு விவரங்களைப் பெற *NSS* ஆய்வுகள் முழுக் கவனத்துடன் பணியாற்றுகின்றன. நாளடைவில் ஒரு சில குறைபாடுகளும் நீக்கப்பட்டு, சிறந்த விவரங்கள் திரட்டப்படும் என்று எதிர்பார்க்கலாம்.

பெறப்படும் விவரங்களில் சில பொருந்தாத (inconsistent) இயல்புகள் இடம்பெற நேரிடுகிறது. ஒரே சீரான வரையறைகள், விதிமுறைகள், செயல்முறைகள் ஆகியவற்றினை உருவாக்கி, நாடு முழுதும் ஒரே மாதிரியான ஆய்வுகள் நடத்த இந்திய அரசு முழு அளவில் முயற்சிகள் மேற்கொள்கிறது. ஆய்வுச் சூழ்நிலைகள் சீராக்கப்பட்டு, சில பொருந்தாத இயல்புகளை நீக்க முயலுதல் இன்றியமையாததாகும்.

தவருன விவரங்கள் அல்லது துல்லியமற்ற விவரங்கள் (inaccurate data) குறிக்கப்படுவது, மிகப் பெரியதொரு குறைபாடாகும். ஆய்வு முடிவுகளை இவை மிகவும் பாதிக்கின்றன. முக்கிய ஆய்வாளர்கள் குறைந்த அளவு ஊதியம் அளிக்கப்படுவதாலோ, அதிகமான வேலைப் பளு தரப்படுவதாலோ, தக்க பயிற்சி பெருததாலோ இத்தகைய குறைபாடு நேரிடலாம். எனவே, இத் துறைகளில் தக்க கவனம் செலுத்தப்படல் வேண்டும். தற்காலத்தில், புள்ளிவிவரச் சேகரிப்புக்கு தனியான அலுவலர்களை அரசினர் நியமித்து, தீவிரமான பயிற்சியும் அளிக்க முற்படுகின்றனர். நாளடைவில் இக் குறையானது நீங்குமென நம்பலாம்.

ஆய்வுப் பணியானது கூட்டுமுயற்சியாக (coordinated) இல்லாதிருப்பின், பல்வேறு அமைப்புகளும் (organisations) இணைந்து செயல்படாத ஒரு நிலையில், ஒரே தன்மையுள்ள விவரங்கள் மீண்டும் சேகரிக்கப்பட்டு, கால விரயமும், பண விரயமும் ஏற்படுகின்றன. எனவே, பல்வேறு துறைகளில் பெறப்படும் விவரங்களை ஒருங்கிணைக்க, இந்திய அரசானது ஒரு தனியான கழகத்தை (*CSO*) உருவாக்கியது. இக் குறையினை நீக்கும்பொருட்டே எனலாம்.

சேகரிக்கப்பட்ட புள்ளி விவரங்களுடன் தக்க முறையில் பகுப்பாய்வு (analysis) நடத்தப்படுவதில்லை என்பதும் விவரங்கள் சரியானபடி செயல்முறைப்படுத்தப் (process) படுவதில்லை என்பதும் அதிகமாகக் கூறப்படும் குற்றச் சாட்டுகளாகும். எனவே, ஆட்சித் துறையின் (நிர்வாகத்தின்) நோக்கத்திற்கு ஏற்ற வகையில் முடிவுகள் பெறப்படும் நிலையானது மாற்றப்பட வேண்டும். பொருளாதார நியமங்களும், நியதிகளும் பெரும்பாலும் நிராகரிக்கப்படுகின்றன.

இந்தியப் புள்ளி விவரங்கள் எளிதில் நோக்கங்களைத் தெளிவுபடுத்தும் தன்மை உடையனவாகக் காணப்படுவதில்லை. எனவே, புள்ளிவிவர ஆய்வுச் செயல்முறைகளும், முடிவுகளும் வளர்ச்சியடைந்த தற்காலத்தில், நாட்டின் பல்வேறு பகுதிகளிலும் (மாநில அளவில்) சேகரிக்கப்படும் புள்ளி விவரங்களின் நோக்கங்களைத் தெளிவுபடுத்த தகுந்த வழிமுறைகள் மேற்கொள்ளப்படவேண்டும்.

புள்ளி விவர ஆய்வு முடிவுகளைப் பொதுமக்களுக்கு அறிவிப்பதில் ஏற்படும் காலதாமதம் (delay) தவிர்க்கப்படவேண்டியது

மிக இன்றியமையாதது. அப்போதுதான், அவை கருதப்பட்ட நோக்கத்திற்குப் பயன்படும் விதத்தில் அமைந்திருக்கும். தற்கால சூழ்நிலையில், ஆய்வுப் பணியானது பல மாநிலங்களிலும் பிரித்துக் கருதப்படும் நிலையல், மக்களுக்குப் பல்வேறு விவரங்களை அறிவிப்பது உடனடியாக இயலும்.

பொதுவாக, இந்திய நாட்டின் பல்வேறு துறைகளிலும் விரிவான அளவில் பெறப்படும் புள்ளி விவரங்களின் ஆய்வுப் பணியானது மிகச் சிறந்த முறையிலும், பாராட்டத்தக்க அளவிலும் அமைந்திருப்பது குறிப்பிடத்தக்க அம்சமாகும். எனினும், இன்னும் அதிக அளவு முன்னேற்றமும், கவனமும், செம்மையான வழிமுறைகளும் தேவைப்படுவன என்று கூறிடின் மிகையாகாது. நாளடைவில் புள்ளிவிவர ஆய்வின் வளர்ச்சி யானது நம் நாட்டில் மிகச் சிறந்து விளங்கும் என்பதில் ஐய மில்லை. உலகிலேயே ஒரு சிறந்த இடத்தைப் பெற்றுத் திகழும் எனவும் கூறலாம்.

நிறைவு

12. வினு வங்கி

(Question Bank)

1. (அ) முழுக் கணிப்பாய்வுடன் ஒப்பிடுகையில், மாதிரி அளவெடுப்புகள் எவ்விதத்தில் சிறந்து பயன்படுகின்றன என விளக்குக.

(ஆ) சென்னை நகரக் கல்லூரி மாணவர் தம் வாழும் சூழ்நிலைகளைப்பற்றி அறிவதற்கு மேற் கொள்ளப்படும் ஒரு மாதிரிஆய்வினை நடத்தும் முறைபற்றிய குத்தாயக் குறிப்பினைத் தருக. [B.Sc., (Anc), A'74]

2. மாதிரியின் (Sampling) அடிப்படைத்தத்துவ (principles) நியதிகள் யாவை ? விளக்கி வரைக.

3. மாதிரிமுறை சார்ந்த பிழை என்றால் என்ன ? தக்க உதாரணத்துடன் விளக்கி வரைக.

4. மாதிரி அளவெடுப்பைத் (sample survey) திட்டமிடுகையில், அமைந்துள்ள பல்வேறு முக்கியக் கட்டங்களை (principal steps) விவரமாகத் தருக.

5. பிறழ்ச்சி என்றால் என்ன என்ன ? அது எவ்வாறு ஏற்படுகிறது ? மாதிரிக் கோட்பாட்டில் (sampling theory) அதன் விளைவுகளை விளக்குக.

6. முன்னோடி ஆய்வு (pilot survey) என்றால் என்ன ? அதன் பயன்கள் யாவை ?

மாதிரிமுறை அளவெடுப்புகளில் மொத்தச் செலவு, மற்றும் மாறுபாடு (variance) இவற்றின் முக்கியத்துவத்தை விவரிக்கவும்.

7. நடைமுறையில் பயன்படுத்தப்பெறும், சில மாதிரி வகைகளை (types of samples) விவரிக்க. ஒன்றோடொன்று ஒப்பிட்டு, அவற்றின் சிறப்புகளையும் அனுகூலங்களையும் ஆய்க.

(B.Sc , A'60)

8. குறிப்பு வரைக : (i) வாய்ப்பு மாதிரித் தேர்வு (probability sampling); (ii) மாதிரி அளவெடுப்புகளில் பாரபட்சம் (பிறழ்ச்சி). (iii) மதிப்பீட்டுத் திட்டம்.

9. (அ) மாதிரி அளவெடுப்பின் பல்வேறு கட்டங்கள் யாவை ?

(ஆ) கீழ்க்காணும் சந்தர்ப்பங்கட்கு எவ்வகை மாதிரித் தேர்வு பொருந்தும் என்று தக்க காரணங்களுடன் கூறுக.

(i) சென்னையில் செய்தித் தாள் படிப்போரின் கருத்துரை அளவெடுப்பு

(ii) தொழிலாளி குடும்பங்களின் குடும்ப பட்ஜெட் கணக்கெடுப்பு. (B.Sc,, (Anc), S'74).

10. மாதிரிக் கணிப்பாய்வின் சிறப்பியல்புகளை விளக்கி, மாதிரி முறைகளில் இயல்நிலைப் பரவலின் பயனை விவரிக்க.

11. பின்வரும் மாதிரித் தேர்வு முறைகளை, உதாரணங் களுடன் விவரிக்கவும் :

(i) சாதாரண ராண்டம் மாதிரி (*SRS*) முறை

(ii) படுகை மாதிரி முறை

(iii) ஒழுங்கு மாதிரி முறை

(iv) பங்கீட்டு மாதிரி முறை

12. ராண்டம் எண்களின் துணையுடன் ஒரு வரம்புற்ற முழுத் தொகுதியிலிருந்து சமவாய்ப்பு மாதிரியினைத் தேர்வு செய்யும் வழிமுறையை விவரிக்க.

13. ஈடுசெய்யப்படாத எளிய ராண்டம் மாதிரி முறையில் (*SRS* without replacement), முழுமைத் தொகுதியின் மாறுபாட்டு (variance) அளவிற்குப் பிறழ்ச்சியற்ற மதிப்பீட்டைப் பெறுக.

14. ஈடுசெய்யப்படாத சாதாரண ராண்டம் மாதிரி முறையில், முழுமைத் தொகுதிச் சராசரி மதிப்பீட்டின் மாதிரி முறை சார்ந்த மாறுபாட்டின் (sampling variance) தொடரினைப் பெறுக.

15. எளிய ராண்டம் மாதிரி முறையில் (திரும்ப வைக்காத முறை),

(i) மாதிரியின் வர்க்கச் சராசரியானது (mean square) முழுமைத் தொகுதியின் வர்க்கச் சராசரியின் பிறழ்ச்சியற்ற மதிப்பீடெனவும்,

(ii) மாதிரி வீதமானது (proportion) முழுமைத் தொகுதி வீதத்தின் பிறழ்ச்சியற்ற மதிப்பீடெனவும் நிறுவுக.

16. மாதிரி அளவினை மதிப்பிடும் பிரச்சினைபற்றி (திட்ட அளவுடன் தொடர்புற்று), பொதுவாக ஆய்க. எளிய ராண்டம் மாதிரி முறையில் எவ்வாறு மாதிரி அளவினைத் தீர்மானிப்பது என விவரிக்க; தீர்வினையும் பெறுக.

17. $\bar{y}$-ல் உள்ள பிழை காரணமாக ஏற்படும் இழப்பு $\lambda \mid \bar{y}-\bar{Y} \mid$ (λ என்பது மாறிலி) எனவும், செலவுச் சார்பு $c = c_a + c_1 \cdot n$ எனவும் தரப்படுகையில், எளிய சமவாய்ப்பு மாதிரி முறையில் (fpc-யினை நீக்கி விட்டு), n-ன் மிகவும் சிக்கனமான மதிப்பு $n = \left(\frac{\lambda S}{c_1 \sqrt{2\pi}} \right)^{2/3}$ என்பதாகக் காட்டுக.

18. N உறுப்புகள் கொண்ட முழுமைத் தொகுதியானது, C, C' என்று இரு பிரிவுகளாகப் பகுக்கப்படுகிறது. சம வாய்ப்பு மாதிரியாக n உறுப்புகளை, இத் தொகுதியினின்று தேர்ந்தெடுக்கையில், C என்ற பிரிவில் இடம்பெறும் உறுப்புகளின் வீதமானது (porportion) p, என்று அமைக்கிறது. இந்த p, மதிப்பானது, முழுமைத் தொகுதி வீதத்தின் (population proportion) பிறழ்ச்சியற்ற மதிப்பீடு என்று காட்டுக. அம்மதிப்பீட்டின் மாறுபாட்டினையும் (variance) தீர்மானிக்க. (B.Sc , A '68)

19. N உறுப்புகள் கொண்ட ஒரு வரம்புடைய தொகுதியிலிருந்து n- அளவுள்ளதொரு எளிய சம வாய்ப்பு மாதிரி தேர்ந்தெடுக்கப்படுகிறது. Np உறுப்புகள் (முழுமைத் தொகுதியில்), ஒரு குறிப்பிட்ட A என்ற தரத்தில் இடம் பெறுகின்றன. m என்பது, தேர்ந்தெடுக்கப்பட்ட மாதிரியில் A தரத்தில் (A grade) உள்ள மொத்த உறுப்புகள் (units) எனில்,

$$E(m) = np \text{ என்றும், } V(m) = \frac{N-n}{(N-1)} \cdot np\,(1-p)$$

என்றும் நிறுவுக. (B.Sc., A. '61)

20. ஒரு சிறந்த மாதிரியின் (good sample) தேவைப்படும் இயல்புகள் யாவை ? எளிய சமவாய்ப்பு மாதிரித்தேர்வு முறையினை விவரிக்கவும். ஓர் எடுத்துக் காட்டினையும் தருக. இத்தகைய மாதிரிமுறையில் பெறப்படும் மாதிரி சராசரியின் எதிர்பார்க்கும் மதிப்பு மற்றும் மாறுபாட்டு அளவு ஆகியவற்றினைப் பெறுக. (B.Sc., A' 67)

21. (அ) எளிய சமவாய்ப்பு மாதிரி என்றால் என்ன ? ராண்டம் எண்கள் பட்டியலைப் பயன்படுத்தி எளிய ராண்டம் மாதிரி ஒன்றினைத் தேர்ந்து எடுக்கும் முறையினை விவரிக்க.

(ஆ) N அளவுள்ள முழுமைத் தொகுதியிலிருந்து, n-அளவுள்ள ஒரு எளிய ராண்டம் மாதிரியானது (ஈடு செய்யப்படாத முறைப்படி) தேர்ந்தெடுக்கப்படுகிறது. மாதிரிச் சராசரியின் மாறுபாட்டு அளவானது $\frac{N-n}{N-1} \cdot \frac{\sigma^2}{n}$ என்பதாகக் காட்டுக. இதில் σ^2 என்பது, முழுமைத் தொகுதியின் மாறுபாடாகும் (variance). (B.Sc., A' 70)

22. இயல்நிலைப்பரவல் அடிப்படையில் அமையும் முழுமைத் தொகுதியிலிருந்து, (சராசரி 50 என்றும், தரவிலக்க மதிப்பு 8 என்றும் தரப்படுகின்றன). ஒரு ராண்டம் மாதிரி தெரிவு செய்யப்படுகிறது. முழுமைத் தொகுதி சராசரியிலிருந்து $\pm$ 0·5 என்ற அளவில் வேறுபடும் மாதிரிச் சராசரி நிலையும் (திட்பம்), 95% நம்பகத் தன்மையும் கொடுக்கப்படின், மாதிரியானது எந்த அளவில் (size) எடுக்கப்படவேண்டும்? (B.Sc. S' 71)

23. எளிய ராண்டம் மாதிரி முறைக்கும், படுகை மாதிரி முறைக்கும் இடையேயுள்ள வேறுபாடுகளை விளக்கி வரைக.

Np வெள்ளைப் பந்துகளும், Nq சிவப்புப் பந்துகளும் உள்ள ஒரு கலனிலிருந்து (urn), n-பந்துகள் மாதியாகத் தேர்ந்தெடுக்கப்பட்டு, p என்ற விகித அளவு மதிப்பிடப்படுகிறது. இம் மதிப்பீட்டின் தரப்பிழையினைத் தீர்மானிக்க. (B.Sc., S' 64)

24. சமவாய்ப்புள்ள படுகை மாதிரி முறையினை விவரிக்க. சமவாய்ப்புள்ள படுகை மாதிரியில், ஒரு தொகுதியின் சராசரிக்கு ஒரு பராபட்சமற்ற (பிறழ்ச்சி) மதிப்பீட்டைக் கண்டு, அம் மதிப்பீட்டின் பரவற்படியின் எண்ணுருக் கோவையை (மாறுபாட்டு அளவு = variance) தருக.

சிக்கலற்ற ராண்டம் மாதிரியைக் காட்டிலும், ஒரு ராண்டம் மாதிரியின் தகுதிகளை ஆய்க. [B Sc. (Anc), A' 74].

25. மாதிரியைத் தேர்ந்தெடுக்கும் முன்னர் ஒரு தொகுதியைப் படுகைகளாக்க வழக்கமான கோட்பாடுகள் யாவை?

படுகைச்சராசரிகளின் வித்தியாசம் (மாறுபாடு), படுகை திட்டவிலக்கங்களின் வித்தியாசம் ஆகிய இரண்டும்தான் சிக்கலற்ற சமவாய்ப்புக் கூறைவிட, குறித்த n கூறளவுள்ள பெரிதும் உகந்த (optimum) பங்கீடு முறையிலான படுகைக் கூறின் லாப தகவுத் திறனுக்குக் (precision) காரணம் என்பதை நிறுவுக.

[B.Sc., (Anc), S' 75]

26. படுகைக் கூறின் (stratified sample) சிறப்புக்கள் யாவை ?

கீழே கொடுக்கப்பட்டுள்ள ஒரு சம வாய்ப்பு உள்ள படுகைக் கூறின் விவரங்களினின்று, தொகுதியின் சராசரியையும் (population mean), அம் மதிப்பீட்டின் பரவற்படியையும் (variance) கண்டுபிடிக்க.

படுகை எண்	அளவு N_i	படுகை பரவற்படி S_i^2	கூறின் அளவு n_i	கூறின் சராசரி $\bar{y}_i$
1	72	44	9	14.2
2	65	27	8	11·6
3	53	35	7	16·1

[B.Sc. (Anc). ,S' 74]

27. ராண்டம் படுகை மாதிரியின் சாதக மற்றும் பாதகங்களை விவரிக்கவும். இம் முறைக்கும், எளிய ராண்டம் மாதிரி முறைக்கும் (*SRS*) இடையில் உள்ள வித்தியாசங்கள் யாவை ?

28. 'வீதப்பங்கீட்டு' (proportional allocation)முறையுடன், படுகை பிரித்தல் (stratification) கையாளப்படுவதை விளக்கி வரைக.

ஒவ்வொரு படுகையிலும் பெறப்படும் மாதிரிச் சராசரியான $\bar{y}_h$ என்பது பிறழ்ச்சியற்று இருக்கும் நிலையில், முழுமைத்தொகுதி சராசரி $\bar{Y}$ க்கு ஒரு பிறழ்ச்சியற்ற மதிப்பீட்டினைக் காண்க.

29. செலவுச் சார்பானது $c = c_0 + \Sigma c_h \cdot n_h$ என்று அமைந்த படுகை மாதிரி (சமவாய்ப்புள்ள) முறையில், n_h அளவானது $N_h \cdot S_h / \sqrt{c_h}$ என்ற மதிப்போடு வீதத் தொடர்புள்ளதாக (proportional to) இருக்கும் நிலையில், $\bar{y}_{st}$ என்ற மதிப்பீட்டுச் சராசரியின் மாதிரிமுறை சார்ந்த மாறுபாடு (sampling variance) சிறும அளவினதாக அமைந்திடும் எனக் காட்டுக.

30. வெவ்வேறு படுகைகளிலும் ஒரு மாதிரியினைப் பங்கிட மேற்கொள்ளப்படும் (i) வீதப் பங்கீட்டு முறை, (ii) உத்தமப் பங்கீட்டு முறை (iii) நேமன் பங்கீட்டு முறை ஆகியவற்றை விவரிக்க.

முழுமைத் தொகுதி சராசரியின் மதிப்பீட்டுக் கோவையினைத் (expression) தருக. அம் மதிப்பீட்டில் மாதிரிமுறை சார்ந்த மாறுபாட்டினை, 'வீதப் பங்கீடு' மற்றும் 'உத்தமப் பங்கீடு' இந் நிலைகளில் ஒப்பிடுக.

31. மாதிரிஅளவு n நிலையானதாக இருக்க, $\frac{n_h}{N_h}$—ஐக்கொண்ட உறுப்புக்கள் (terms) பொருட்படுத்தப்படாது (நீக்கத்தக்க அளவில்) இருக்கும் நிலையில், $V_{opt} \leq V_{prop} \leq V_{ran}$ என்று நிரூபிக்க. இதில் 'உத்தமப்' பங்கீடானது நிலையான n அளவிற்கு அமைவதாகும் (அதாவது, n_h என்பது $N_h \cdot S_h$ உடன் வீதத் தொடர்பு உடையதாகும்.)

32. படுகை மாதிரி முறையில், மதிப்பிட்ட சராசரியானது V_0 என்ற நிலையான மாறுபாட்டு (variance) அளவினை உடையதாக இருக்கவேண்டுமெனில், (i) வீதப் பங்கீடு மற்றும் (ii) உத்தமப் பங்கீடு இவ்வழி முறைகளில், n என்ற மாதிரி அளவு எவ்வாறு அமையும் என்பதைப் பெறுக. (n-ன் கோவைகளை எழுதுக.)

33. முன்னோடியாக, S_h மதிப்பை மதிப்பிடுவதற்கு, n_0 என்ற மாதிரி அளவு கொள்ளப்படுவதாக வைத்துக்கொண்டு, நேமன் பங்கீட்டு முறையானது வீதப் பங்கீட்டு முறையினைவிட அதிகத் திட்பமுள்ள மதிப்பீட்டைத் தரவேண்டுமெனில், முதலில் தெரிவுசெய்யப்படும் மாதிரியானது எவ்வளவு பெரிய அளவினதாக இருக்கவேண்டும் என்பதைக் காண்க.

34. ஒரு தொகுதியானது N_1, N_2 அளவுகள் உள்ள இரு படுகைகளாகப் பிரிக்கப்பட்டு, $S_1{}^2$, $S_2{}^2$ என்ற மாறுபாடுகளுடன் (variances) (அல்லது பரவற்படி) காணப்படுகிறது. இரு படுகைகளின்றும் முறையே n_1, n_2 அளவுள்ள சமவாய்ப்பு மாதிரிகள், திரும்பவைக்காத முறைப்படி தேர்ந்து எடுக்கப்படுகின்றன. முழுமைத் தொகுதி சராசரியின் மதிப்பீட்டையும், அம் மதிப்பீட்டின் மாறுபாட்டு அளவையும் (பரவற்படி) கணக்கிடுக. (B.Sc., A'70)

35. தொகுதிப் படுகைகள் :

படுகை I : 1, 0, 1

படுகை II : 1, 0, 0, 1

என்று தரப்படுகின்றன. இவற்றிலிருந்து $n_1 = 2$; $n_2 = 3$ என்ற அளவுகளில் எல்லா மாதிரிகளேயும் தேர்ந்தெடுத்து. $E(pst) = P$ என்று நிரூபிக்க.

36. (i) வீதப்பங்கீடு, (ii) உத்தமப்பங்கீடு ஆகிய வழிமுறைகளில் $n = 5$ என்ற அளவுள்ள மாதிரியினைப் பங்கீடு செய்க. முழுமைத் தொகுதி :

படுகை 1 : 2, 5, 8, 9

படுகை 2 : 9, 10, 12, 15, 15, 16

என்று தரப்படுகிறது. ($c_1 = 2$; $c_2 = 4$ எனக் கொள்க.)

37. N அளவுள்ள முழுமைத் தொகுதியானது, N_1, N_2 என்ற அளவுகளுள்ள இரு படுகைகளாகப் பிரிக்கப்படுகிறது. n அளவுள்ள ஒரு படுகை மாதிரியின் அடிப்படையில், முழுமைத் தொகுதிச் சராசரியின் பிறழ்ச்சியற்ற மதிப்பீட்டைப் பெறுக. விகித அளவில் (proportionate) மாதிரித் தேர்வு நடைபெறுகையில், இம்மதிப்பீட்டின் திட்டபிழையினைத் (standard error) தீர்மானிக்க. (B.Sc., A' 69)

38, n அளவுள்ள மாதிரியினை, ஒரு முழுமைத் தொகுதியின் பல படுகைகட்கும், மதிப்பீட்டுச் சராசரியின் மாறுபாட்டு அளவு (பரவற்படி) சிறும மதிப்பினதாக ஆகும் வகையில் பங்கிடப்பட வேண்டும் எனில், (i) i ஆவது படுகையில் உள்ள மாதிரி அளவு n_i மற்றும் (ii) மதிப்பீட்டுசராசரியின் சிறும மாறுபாடு (minimum variance) ஆகியவற்றைப் பெறுக.

39. வீத மதிப்பீட்டைக் கருதி (estimation of proportion) படுகை மாதிரி முறையில்

$$V_{ran} = V_{prop} + \frac{\Sigma w_h (P_h - P)^2}{n};$$

$$V_{prop} = V_{opt} + \frac{\Sigma w_h (\sqrt{P_h Q_h} - \sqrt{\overline{P_h Q_h}})}{n}$$

என்று நிரூபிக்க இவற்றில்,

$$\sqrt{\overline{P_h Q_h}} = \Sigma\ w_h \sqrt{P_h Q_h} \text{ ஆகும்.}$$

$\left(w_h = \frac{N_h}{N}\right.$ என்று கொள்க).

40. செலவுச் சார்பு $c = c_0 + \Sigma t_h \cdot \sqrt{n_h}$ என்று இருக்கையில், (c_0, t_h ஆகியவை கொடுக்கப்பட்ட எண்கள்),

$V(\bar{y}_{st})$-யினைச் சிறும அளவினதாக்கும் வகையில் அமையும் n_h-ஆனது (நிலையான மொத்தச் செலவுடன்), $\left(\frac{W_h^2 \cdot S_h^2}{t_h}\right)^{2/3}$ என்ற மதிப்புடன் வீதத் தொடர்பு பெற்றிருக்கும் எனக் காட்டுக.

1000 என்ற மாதிரி அளவினைக் கொண்டு பின்வரும் நிலையில், n_h அளவுகளைத் தீர்மானிக்க.

படுகை	w_h	S_h	t_h
1	0·4	4	1
2	0·3	5	2
3	0·2	6	4

41. ஒழுங்கு மாதிரி முறையினை விவரிக்க. இம் முறையினை, மதிப்பீட்டு மாறுபாடுகளைப் (பரவற்படிகளை) பொறுத்த அளவில், படுகை மாதிரி முறையுடனும்) எளிய சமவாய்ப்பு மாதிரி முறையுடனும் ஒப்பிடுக.

42. ஒழுங்கு மாதிரி எவ்வாறு தேர்ந்தெடுக்கப்படுகிறது? எத்தகைய அளவெடுப்புகளில் இம்முறையானது நடைமுறையில் பயன்படுத்தத்தக்கதாகும் ?

சில கருதுகோள்களுடன் (assumptions) (எவை எனக் குறிப்பிட்டு), ஒழுங்கு மாதிரி முறையானது, K பெரிய கூறுகளைக் கொண்ட ஒரு தொகுதிக்கு மேற்கொள்ளப்படும் எளிய சமவாய்ப்பு முறையாகவே அமைவதைக் காட்டுக.

43. முழுமைத் தொகுதி சராசரியின் பிறழ்ச்சியற்ற மதிப்பீட்டையும், அம் மதிப்பீட்டின் மாதிரிமுறை சார்ந்த மாறுபாட்டையும், ஒழுங்கு மாதிரியில் கணக்கிட்டுப் பெறுக.

44. ஒழுங்கு மாதிரி முறையில், 'காலச் சுழற்சி' periodicity) உள்ள தொகுதிகள் (காலவட்ட ஒழுங்குள்ள தொகுதிகள்) அமையப்பெறும் ஒரு நிலையினை விளக்குக. அதனால், மாதிரித் தேர்வு எவ்வாறு பாதிக்கப்படும் என உதாரணத்துடன் விவரிக்க.

45. (அ) ராண்டம் படுகைக் கூறில் விகிதப்படி பிரிக்கும் பங்கு முறை, பெரிதும் உகந்த (உத்தமப்) பங்கு முறை என்றால் என்ன ? அவற்றின் சிறப்புகள் யாவை?

(ஆ) முறையாகக் கூறு எடுத்தல் முறையினை (systematic) திறனாய்வு செய்க. [B.Sc., (Anc) S 75.]

46. $N = nK$ என்று அமையும் முழுத் தொகுதியிலிருந்து, மாதிரித் தேர்வு இடைவெளியினை K எனக்கொண்டு (sampling interval K), n-அளவுடன் எடுக்கப்படும் எல்லாவித ஒழுங்கு மாதிரிகளையும் முறையாகப் பட்டியல் அமைத்துக் காட்டுக.

ஒரு ஒழுங்கு மாதிரியின் சராசரியானது முழுமைத் தொகுதிச் சராசரியின் பிறழ்ச்சியற்ற மதிப்பீடெனக் காட்டுக. இம் மதிப்பீட்டின் மாறுபாடு $\left(\frac{K-1}{K}\right) S_c^2$ என்றும் காட்டுக.

[(K என்பது மாதிரித் தேர்வு இடைவெளி ; S_c^2 என்பது மாதிரிச் சராசரிகளின் சராசரி வர்க்கம் (mean square) என்று கொள்க. [B.Sc. (Revised) S'75]

47. (அ) படுகை மாதிரி முறையினை விளக்கி அதன் பயனை ஓர் எடுத்துக்காட்டின்மூலம் விளக்குக. படுகை மாதிரி முறையில் 'உத்தமப் பங்கீட்டு'ப் பிரச்னையினை ஆராய்க.

(ஆ) ஒழுங்கு மாதிரி முறையினை விவரிக்க.

[B.Sc., (old) S' 75]

48. ஒழுங்கு மாதிரி முறையில்

$$V(\bar{y}_{sys}) = \frac{S^2}{n} \quad \frac{N-1}{n} \quad (1 + \overline{n-1}.\rho_w)$$

என நிரூபிக்க. ρ_w என்பது பிரிவு உள்ளீட்டு ஒட்டுறவாக (intra class correlation) வரையறுக்கப்படுவதாகும்.

49. திரள் மாதிரி முறையினை, இருகட்ட மாதிரி அடிப்படையில் விவரிக்கவும். முழுமைத்தொகுதியின் மொத்தம் (total) மற்றும் சராசரி இவற்றின் திரள் மாதிரி முறை மதிப்பீடுகளைப் பெறுக.

$E(\hat{Y}) = Y$ என்று நிரூபிக்க.

50. திரள் மாதிரித் தேர்வினை, எளிய சமவாய்ப்பு மாதிரித் தேர்வு முறையுடன், மதிப்பீட்டு மாறுபாட்டைப் பொறுத்த அளவில், ஒப்பீடு செய்க. அதில் உள்ள சில குறிப்பிட்ட நிலைகளையும் விளக்குக.

51. மாதிரிப் பங்கீட்டுப் பிரச்னை திரள் மாதிரித் (இருகட்ட) தேர்வில் எவ்வாறு அமைகிறதென விளக்குக.

$C = c_1 . m + c_2 . mn^1$. என்ற செலவுச் சார்புடன், உத்தமப் பங்கீட்டு முறைப்படி, n', m ஆகியவை எவ்வாறு மதிப்பிடப்படுகின்றன ? அவற்றின் மதிப்புகளையும் பெறுக.

52. முழுத் தொகுதி வீதத்திற்கு (Proportion) திரள் மாதிரி (இருகட்ட) முறையில் மாதிரி மதிப்பீடு என்ன என்று கணக்கிடுக. $E(\hat{P}) = P$ என்று நிறுவுக. அம்மதிப்பீட்டின் மாறுபாட்டு மதிப்பையும் (sampling variance) நிர்ணயிக்க.

53. திரள் மாதிரிமுறை அடிப்படையில், $\hat{Y}$-யின் மாதிரி முறை சார்ந்த மாறுபாட்டின் தொடரை (expression) பெறுக. அதன் மாதிரி மதிப்பீடு என்ன என்று எழுதுக. அம்மதிப்பீடானது பிறழ்ச்சியற்றதென நிறுவுக.

54. பலகட்ட மாதிரி அளவெடுப்பு முறை, மற்றும் பல் தோற்ற மாதிரி அளவெடுப்பு முறை (multi-phase) இவ்விரு முறைகளையும் வேறுபடுத்தி, விரிவான முறையில் ஆராய்க தக்க உதாரணங்களும் தருக.

55. ‘அளவோடு விகித அளவில் மாறுபடும் வாய்ப்பு' அல்லது PPS மாதிரி முறைத் தேர்வு என்றால் என்ன? ‘எளிய' திரள் முறையிலிருந்து இது எவ்விதம் வேறுபடுகிறது?

56. PPS மாதிரி முறையில், முழுமைத் தொகுதி மொத்தத்தின் மதிப்பீட்டைப்பெற்று, அதன்மூலம் $\hat{Y}$ என்ன என்று மதிப்பிடுக. இம்மதிப்பீடு பிறழ்ச்சி உடையதா என ஆராய்க.

57. PPS மாதிரி முறையில், மாதிரி முறையைச் சார்ந்த மாறுபாட்டின் தொடரை எழுதுக. அதன் மாதிரி மதிப்பீடு என்ன?

மேலும், $E[\hat{V}(\bar{y}_{pps})] = V(\bar{y}_{pps})$ என நிறுவுக.

58. PPS மாதிரி முறையில், மாதிரிப் பங்கீடு எவ்வாறு அமைகிறது என ஆராய்க. இறுதிநிலைத் திரள்கள் அனைத்தும் ஒரே அளவினதாக $(n_s = n')$ தெரிவு செய்யப்படுவதாகக் கொள்க. செலவுச் சார்பின் $C = c_1 m + c_2 \cdot mn'$ என்றும் எடுத்துக் கொண்டு ஆய்க. சில நிலைகளை விளக்குக.

59. எளிய திரள் முறையுடன் (இரு கட்டத் தேர்வில்), *PPS* மாதிரி முறையினை, அவற்றின் மாறுபாடுகள் வாயிலாக, ஒப்பீட்டு முடிவுகளைப் பெறுக.

60. 'வீத மதிப்பீடு' (proportions) பற்றிய *PPS* மாதிரி முறையினை விளக்கி $\hat{P}_{pps}$ மற்றும் $\hat{V}(\hat{P}_{pps})$ ஆகியவற்றின் தொடர்களை, வழக்கமான அடிப்படையில் பெறுக.

61. விகித மதிப்பீட்டு முறையினை விளக்குக. சில எடுத்துக் காட்டுகளைத் தருக. இம் முறையானது மற்ற மாதிரி மதிப்பீட்டு முறைகளினின்று எவ்வாறு வேறுபடுகிறது ?

62. முழுமைத் தொகுதி விகிதமான R-ஐக் கருதி, $\hat{R}$ மற்றும் $V(\hat{R})$ ஆகியவற்றின் பொதுவான தொடர்களைப் பெறுக. $\hat{R}$ ஆனது பிறழ்ச்சியற்றதா என ஆராய்க.

63. (அ) எளிய ராண்டம் முறையில் மாதிரி விகிதத்தினை (sample-ratio) வரையறை செய்து, அதன் பிறழ்ச்சி மற்றும் மாறுபாடு ஆகியவற்றின் தொடர்களைப் பெறுக.

(ஆ) r என்ற மாதிரி விகித மதிப்பானது எப்போது பிறச்ச்சியற்று அமைகிறதென ஆராய்க.

64. விகித மதிப்பீட்டு முறையினை, ஒரு மாறியுடன் அமையும் எளிய ராண்டம் மாதிரி முறையுடன் ஒப்பிடுக.

65. விகித மதிப்பீட்டு முறையினை, படுகை மாதிரித் தேர்வில் பயன்படுத்தி, (i) இணைத்த முறை மற்றும் (ii) தனித்த முறை ஆகியவற்றை விளக்கி, இரண்டிலும் மாதிரி விகித மதிப்பு (r), மற்றும் $V(r)$ ஆகியவற்றைப் பெறுக.

66. தொடர்பு மதிப்பீட்டினை (regression estimate) (i) சாதாரண ராண்டம்முறை; (ii) படுகை மாதிரி முறை ஆகிய வற்றில், விளக்கமாக ஆராய்க.

67. மாதிரி முறை சார்பற்ற பிழைகள் யாவை ? அவை ஏற் படுவதற்குரிய காரணங்களைத் தொகுத்துக் கூறுக. இத்தகைய பிழைகளைக் குறைப்பதற்கு எடுக்கவேண்டிய சில நடவடிக்கை களைக் குறிப்பிடுக. பெரிய அளவில் அமையும் ஆய்வுகளைப் பொறுத்து, இவற்றை விளக்குக.

68. மாதிரி முறை சார்பற்ற பிறழ்ச்சி, மற்றும் மாறுபாடு இவற்றைக் கணக்கிடும் முறையினை விளக்குக. எளிய சமவாய்ப்பு

முறையில் இவற்றின் தொடர்களை (expressions), வழக்கமான முறையில் பெறுக.

69. பதிலின்மை (non-response) மூலம் பிழைகள் எவ்வாறு நிகழ்கின்றன ? அவற்றைப்பற்றிய சில கணக்கியல் முடிவுகளைப் பெறுக.

70. மாதிரி முறை சார்பற்ற பிழைகளைக் கட்டுப்படுத்தும் சில வழிமுறைகளைக் குறிப்பிட்டு, விரிவாய்வு செய்க.

71. பதிலளிப்பில் ஏற்படும் பிழைகளைப்பற்றி விரிவாக ஆராய்க. அவற்றை எவ்வாறு மட்டுப்படுத்துவது என்பதற்கான சில வழிமுறைகளைக் குறிப்பிடுக.

72. ஒருமாதிரிக் கணிப்பின் முக்கிய இயல்புகள் (feature) யாவை ? மாதிரித் தேர்வு முறைகளில், இயல்நிலைப் பரவல் எவ்வாறு பயன்படுகிறதென்று விளக்கி வரைக. (B. Sc., S' 75)

73. பெரிய அளவிலான (large scale) மாதிரி அளவெடுப்புகளின் (surveys) பொதுத் தன்மைகளை விவரிக்க.

இந்தியாவில் நடத்தப்படும் தேசீய மாதிரி அளவெடுப்புகளின் குறிக்கோள்களை, அளவெடுப்புக்கள் ஆராய்ந்த சமுதாய, பொருளாதார, மக்கள் வாழ்வு சம்பந்தப்பட்ட இனங்களை (factors) வெளிக் கொணர்ந்து ஆராய்க. [B.Sc., (A.n.,) S'75]

74. நம் நாட்டில் தேசிய மாதிரி ஆய்வு எந்த விதத்தில் நடைபெறுகிறது என்று விளக்கி வரைக.

75. மக்கட் கணிப்பின் முக்கியத்துவத்தினை நன்கு விவரித்து இந்திய மக்கட் கணிப்பானது எவ்வாறு நடைபெறுகிறது என விளக்குக. இவ்வாய்வானது மேலும் செம்மைப்படுத்தப்பட சில நடவடிக்கைகளைக் குறிப்பிடுக.

76. 1961ல் நடைபெற்ற மக்கட் கணிப்பாய்வினைப்பற்றி விரிவாக ஆய்க.

77. இந்திய மக்கள் கணிப்பின் முக்கிய அம்சங்கள் யாவை? 1971ல் மக்கள் கணிப்பாய்வு எந்த விதத்தில் நடைபெற்றது என்று விளக்கி வரைக. அதில் பயன்படுத்தப்பட்ட வினாப் பட்டியலில் இடம் பெற்ற விவரங்களைப் பற்றிய விளக்கமும் தருக.

78. இந்திய நாட்டின் பிறப்பு இறப்பு விவரங்கள் எவ்வாறு அதிகார பூர்வமாகப் பெறப்படுகின்றன என்று குறிப்பிடுக.

79. இந்திய மக்கட் கணிப்பு ஆய்வுகள் பத்து ஆண்டு கட்கு ஒருமுறை எவ்வாறு நடைபெறுகின்றன என்று விவரிக்க வும். 1971-ல் நடைபெற்ற இந்திய மக்கட் தொகைக் கணிப்பின் குறிப்பிடத்தக்க அம்சங்களைத் தருக. (B.Sc., A'73)

80. இந்திய நாட்டின் புள்ளி விவரங்களின் ஆய்வுபற்றிய ஒரு திறனாய்வு செய்க. C. S. O. அமைப்பின் தன்மையையும், முக்கியப் பணிகளையும் விவரிக்க.

81. (அ) மக்கட் கணிப்பில் பொதுவாகக் காணப்படும் தத்துவக் கோட்பாடுகளை (principles) விரிவாக ஆய்க. நம் நாட்டில் பயன்படுத்தப்படும் வழிமுறைகளைப்பற்றி விவரிக்க.

(ஆ) 1971 இந்திய மக்கட் கணிப்பில் பெறப்பட்ட முக்கிய முடிவுகளைத் தொகுத்து வரைக. (B.Sc., S'75)

82. இந்தியப் புள்ளி விவர ஆய்வுகள்பற்றிய பொதுவான திறனாய்வினை அளிக்கவும்.

83. இந்திய விவசாயப் புள்ளி விவரங்கள் சேகரிக்கப்படும் முறையை விவரிக்க. குறைபாடுகளையும், அவற்றைக் களைய வழிமுறைகளையும் தொகுத்தளிக்க.

84. நம் நாட்டின் தொழில் துறைப் புள்ளி விவர ஆய்வு நோக்கங்களையும், தன்மையையும் விவரமாக வரைக.

85. இந்திய நாட்டின், வர்த்தக நிலை பற்றிய புள்ளி விவர ஆய்வுகளை விவரமாக விளக்கி வரைக. குறியீடுகள் பெறப்படும் விதத்தையும் விளக்குக.

வினா வங்கி (இணைப்பு)

1. 'நிகழ்தகவு மாதிரி' யினை (probability sample) வரையறை செய்க. இரு உதாரணங்கள் தருக.

ஈடு செய்யப்படாத (without replacement) முறையில், 'n' அளவுள்ள ஒரு ராண்டம் மாதிரியானது N அளவுள்ள முழுத் தொகுதியிலிருந்து தெரிவுசெய்யப்படுகிறது. முழுமைத் தொகுதி சராசரியின் பிறழ்ச்சியற்ற மதிப்பீட்டையும், மதிப்பீட்டின் திட்டப் பிழையினையும் (standard error) பெறுக.

மாதிரித் தெரிவானது, ஈடுசெய்யப்படும் (with replacement) முறையில் அமைகையில், முழுமைத் தொகுதி மொத்தத்தின் மதிப்பீட்டினுடைய திட்டப் பிழையானது என்னவாக இருக்கும்?

(B Sc., A. '76)

2. எளிய ராண்டம் மாதிரி முறையுடன் ஒப்பிடுகையில், படுகைப்படுத்திய ராண்டம் மாதிரி முறையில் உள்ள அனுகூலங்கள் யாவை?

ஒரு முழுமைத் தொகுதியானது மூன்று படுகைகளாகப் பிரிக்கப்படுகிறது.

(i) வீதப் பங்கீடு (ii) உத்தமப் பங்கீடு (நிலையான மாதிரி அளவைக் கொண்டு) ஆகிய இரு முறைகளில் இரு தனித்த மாதிரிகள் தெரிவு செய்யப்படுகின்றன. முழுமைத் தொகுதிச் சராசரிக்கு இம்மாதிரிகளினின்று பெறப்படும் மதிப்பீடு களின் மாறுபாடுகளை ஒப்பிடுக. (B.Sc., A. '76)

3. இந்திய அரசின் புள்ளி விவரங்களின் மூலக்கூறுகளைப் பற்றிய ஒரு திறனாய்வு செய்க. குறிப்பாக, விவசாய மற்றும் தொழில் துறைகளை மேற்கொள்க. (B.Sc., A. '76)

மேற்கோள் நூற்பட்டியல்
(Bibliography)

1. Sample Survey Methods and Theory (Vol. I & Vol. II) —*M. H. Hansen, W. N. Hurwitz and W. G. Madow.*
2. Sampling Techniques—*W. G. Cochran.*
3. Elementary Sampling Theory—*Taro Yamane.*
4. Sampling Theory and Methods—*M. N. Murthy.*
5. The Design of Sample Surveys—*Des Raj.*
6. Sampling Methods for Censuses and Surveys—*F. Yates.*
7. Sampling Theory of Surveys with Applications —*P. V. Sukhatme.*
8. Some Theory of Sampling—*W. E. Deming.*
9. Sample Design in Business Research—*W.E. Deming.*
10. Elements of Statistics—*B. N. Asthana.*
11. Fundamentals of Statistics—*D. N. Elhance.*

கலைச்சொற்கள்

A

Absolute	— தனியான
deviation	— விலக்கம்
value	— பெறுமானம், மதிப்பு
Abstract	— புலனுகாத
Accuracy	— பிழையின்மை, துல்லியம்
Accurate	— துல்லியமான
Add	— கூட்டு
Additive	— கூட்டுத்தன்மை
law	— விதி
Aggregate	— மொத்தம்
Agricultural statistics	— விவசாயப் புள்ளிவிவரம், வேளாண்மைப் புள்ளிவிவரம்
Allocation	— பங்கீடு
of sample (see stratified, cluster, sampling ratio estimate)	— மாதிரிப் பங்கீடு
Analysis	— பகுப்பாய்வு
of variance	— பரவற்படிப் பகுப்பாய்வு, மாறுபாட்டுப் பகுப்பாய்வு
Analytic method	— பகுப்பாய்வு முறை
Approximate	— தோராயமாக
Applied statistics	— பயன்முறைப் புள்ளியியல்
Area sampling	— பரப்பு கருதிய மாதிரித் தேர்வு (புலன்வழி மாதிரித் தேர்வு)
Ascending order	— ஏறுவரிசை
Ascertainment error	— உறுதிசெய்தலிற் பிழை
Autocorrelated	— தற்தொடர்புடைய
Auxiliary information (see Ratio Estimate Regression estimate)	— துணைவிவரம்
Average	— சராசரி

B

Base — அடிப்படை
  year — ஆண்டு
Best fit — உத்தமப் பொருத்தமான
Bias — பிறழ்ச்சி, ஒருபுறச் சாய்வு
Binomial distribution — ஈருறுப்புப் பரவல்
Birth-rate — பிறப்புவீதம்
Bivariate — இருமாறி
Blank — வெற்றிடம்
Block — தொகுதி
Boundary — எல்லை, வரம்பு
Business cycle — தொழிற்சுழற்சி

C

Calculate — கணக்கிடு
Calculator — கணக்கீட்டுப்பொறி
Calculus — நுண்கணிதம்
  differential — வகை நுண்கணிதம்
  integral — தொகை நுண்கணிதம்
Cancel — நீக்கு
Capital — முதலீடு, மூலதனம்.
Case — நிலை
  general — பொது நிலை
  particular — குறிப்பிட்ட நிலை
Census — முழுக் கணிப்பு
  population — மக்கட்கணிப்பு
  method — முறை
  report — அறிக்கை
  return — விவரப்பட்டியல்
Central Statistical Organisation — மத்தியப் புள்ளிவிவர ஆய்வுக் கழகம்
Chain base index — சங்கிலிக் குறியீட்டெண்
Chance — வாய்ப்பு
Characterisitic — சிறப்பியல்பு
Check — சரிபார்த்தல், தணிக்கை
Class interval — பிரிவு இடைவெளி
Classification — பாகுபாடு, பிரிவினை.
Classify — பிரிவுபடுத்தல்
Cluster — திரள்
Coefficient — கெழு
  of correlation — ஒட்டுறவுக் கெழு

Coefficient of dispersion — சிதறல் கெழு
  of variation — மாறுபாட்டுக் கெழு
Complete count — முழுக் கணிப்பு
Commodity — பொருள்
Comprehensive — அகலவிரிவான
Computers — கணிகள்
  electronic — மின்னுக்கக் கணிகள்
Combined — இணைத்த
  ratio estimate — விகித மதிப்பீடு
  regression estimate — தொடர்பு மதிப்பீடு
Concentration — குவிவு
Confidence — நம்பகத் (தன்மை)
  coefficient — கெழு
  interval — இடைவெளி
  limits — எல்லைகள்
Consistent — பொருத்தமான
  estimator — மதிப்பீட்டு அளவை
Consistency — பொருத்தம்
Constant — மாறிலி
Contingency table — இணைப்புப் பட்டியல்
Continuous variable — தொடர்மாறி
Correction — திருத்தம்
Correlation — ஒட்டுறவு, உடன்தொடர்பு
  linear — நேர்கோட்டு ஒட்டுறவு
  intraclass — பிரிவுகள் உள்ளிட்ட ஒட்டுறவு
  negative — எதிர்கணிய ஒட்டுறவு
  partial — பகுதி ஒட்டுறவு
  positive — நேர்கணிய ஒட்டுறவு
Cost — செலவு
  function — சார்பு
Covariance — உடன் மாறுபாடு
Covariation — உடன்மாற்றம்
Cumulative frequency — குவிவு அலைவெண்
Cycle — சுழற்சி

D

Data — விவரங்கள்
  observed — கண்டறிந்த விவரங்கள்
  primary — முதனிலை விவரங்கள்

Data secondary — வழிநிலை விவரங்கள்
statistical — புள்ளி விவரங்கள்
Deathrate — இறப்புவீதம்
Decimal — தசம
Decennial — 10 ஆண்டுகட்கு ஒரு முறை
Define — வரையறு
Definition — வரையறை
Degrees of freedom — வரையற்ற பாகைகள் (கூறுகள்)
Denominator — விகுதி
Density — அடர்த்தி
,probability — நிகழ்தகவு அடர்த்தி
Dependent — சார்ந்த
variable — மாறி
Descriptive — விளக்கமான
Design — வடிவம், உரு
,Sample — மாதிரித் திட்டம், உருவமைப்பு
Demography (see Vital Statistics) — மக்களியல்
Deviation — விலக்கம், வேறுபாடு
absolute — தனிவிலக்கம்
standard — தர விலக்கம்
mean square — சராசரி வர்க்க விலக்கம்
Diagram — விளக்கப்படம்
Digit — இலக்கம்
Dispersion — சிதறல்
Distribution — பரவல்
Binomial — ஈருறுப்புப் பரவல்
continuous — தொடர் பரவல்
frequency — அலைவுப் பரவல்
normal — இயல்நிலைப் பரவல்
poisson — பாய்ஸான் பரவல்
probability — நிகழ்தகவுப் பரவல்
Division — பிரித்தல், பகுத்தல்

E

Economic Statistics — பொருளாதாரப் புள்ளிவிவரம்
Effective — பயனுடைய, சக்திவாய்ந்த
Efficiency — திறன்
Elementary — தொடக்கநிலை

Enquiry	—	கணக்கீடு, கணக்கெடுப்பு
Enumeration	—	கணிப்பு
Enumerator	—	கணிப்பாளர்
End correction	—	எல்லைத் திருத்தம்
Error	—	பிழை, வழு
,experimental	—	செய்முறைப் பிழை
,non-sampling	—	மாதிரி முறைசாராத பிழை
,relative	—	சார் பிழை
,sampling	—	மாதிரி முறைசார்ந்த பிழை
,standard	—	தரப் பிழை
Estimate	—	மதிப்பிடு, மதிப்பீடு
Estimator	—	மதிப்பீட்டு அளவை
biased	—	பிறழ்ச்சியுடைய மதிப்பீட்டு அளவை
consistant	—	பொருத்தமானமதிப்பீட்டு ,,
unbiased	—	பிறழ்ச்சியற்ற மதிப்பீட்டு ,,
Exchange	—	பரிவர்த்தனை
Expected value	—	எதிர்பார்க்கும் மதிப்பு
Expectation of life	—	எதிர்பார்க்கும் ஆயுள்
Expression	—	தொடர், கோவை

F

Factor	—	காரணி
Field Enquiry	—	கள ஆய்வு
Field work	—	களப் பணி
Finite	—	வரம்புடைய
Population	—	முழுத்தொகுதி
Population Correction	—	திருத்தம்
Fluctuation	—	ஏற்ற இறக்கம்
seasonal	—	பருவகால ஏற்ற இறக்கம்
random	—	ராண்டம் ஏற்ற இறக்கம்
Frame	—	‘ சட்டம் ’
sampling	—	மாதிரிச் ‘ சட்டம் ’
Frequency	—	அலைவெண்
distribution	—	பரவல்
curve	—	வளைவு
table	—	பட்டியல்
Function	—	சார்பு

G

Geographic	—	பூகோள, புவியியல் தொடர்புடைய
classification	—	பாகுபாடு

Grade — தரம்
Graph — வரைபடம்
Group — பிரிவு
Grouping — தொகுப்பு

H

Heterogeneous — பலபடித்தான
Homogeneous — ஓரின, ஒருபடித்தான
Hypothesis — எடுகோள்
,test of — சோதனை

I

Identity — சர்வசமம், முற்றொருமை
Imaginary — கற்பனை
population — முழுத்தொகுதி
Inaccurate — செம்மையற்ற
Inconsistent — பொருந்தாத
Independent — தனித்த
Indicator — குறிப்பீடு
Indian Statistics — இந்தியப் புள்ளிவிவரங்கள்
Indian Statistical Institute — இந்திய ஆய்வு நிறுவனம்
Individual — தனித்த, தனிநபர்
Indian Council of Agricultural Research — இந்திய விவசாய ஆராய்ச்சிக் கழகம்
Inequality — சமனிலி
Inertia — மாறுத் தன்மை
of large numbers — பேரினங்களின் மாறுத் தன்மை
Infinite — வரம்பற்ற
Infinity — எண்ணிலி
Inference — உய்த்துணர்வு, முடிவு
Initial — தொடக்கத்திலுள்ள
Input — முதற்பொருள், உட்பாடு
Interaction — இடைவினைவு
Interval — இடைவெளி
class — பிரிவு இடைவெளி
confidence — நம்பக இடைவெளி
Investigator — ஆய்வாளர்
Investigation — ஆய்வு
field — கள ஆய்வு
statistical — புள்ளிவிவர ஆய்வு
Item — உறுப்பு, இனம்

J

Joint	—	இணைந்த
Judgement Sampling	—	தீர்ப்பு மாதிரித் தேர்வு
Journals	—	பத்திரிகைகள்

L

Large sample	—	பெரியஅளவு மாதிரி, விரிந்த மாதிரி
List	—	பட்டியல்
household	—	குடும்பப் பட்டியல்
houses	—	வீடுகள் பட்டியல்
frame	—	'சட்ட'ப் பட்டியல்
Law of Statistical Regularity	—	புள்ளிவிவர ஒழுங்கு விதி
Least squares method	—	குறைந்த வர்க்க முறை
Limit	—	எல்லை, வரம்பு
Living standard	—	வாழ்க்கைத் தரம்
Lower confidence limit	—	கீழ் நம்பிக்கை எல்லை
Loss	—	இழப்பு
Logical	—	தர்க்கரீதியான
Linear	—	நேர்கோட்டு
trend	—	போக்கு

M

Machine	—	பொறி
Mail enquiry	—	அஞ்சல்வழி ஆய்வு
Manufacturing	—	உற்பத்தி, ஆக்கம்
Map	—	படம்
Mathematical	—	கணக்கியல்
Maximum	—	மீப்பெரு, பெரும, உச்சம்
likelihood	—	உச்சநிகழ் வாய்ப்பு
Mean	—	சராசரி
,weighted	—	நிறையிட்ட சராசரி
square error	—	சராசரி வர்க்க விலக்கம் (பிழை)
Method	—	முறை
Minimum	—	சிறும, மீச்சிறு
variance	—	மாறுபாடு
Mortality	—	இறப்பு
rate	—	வீதம்
Most probable	—	பெரும்பாலும் நிகழக்கூடிய
Mutually exclusive	—	ஒன்றையொன்று விலக்கத்தக்க
Multivariate	—	பல்மாறி

Multiphase — பன்முக
Multistage — பலகட்ட
sampling — மாதிரித் தேர்வு
Multi subject — பல்பொருள்

N

National — தேசிய
Sample Survey — மாதிரி ஆய்வு (அளவெடுப்பு)
Natural resources — இயற்கை வளம்
Nearly — ஏறத்தாழ
Negligible — தவிர்க்கத்தக்க, நீக்கத்தக்க
Net income — நிகர வருமானம்
value — மதிப்பு
Neyman Allocation — மாதிரிமுறை நேமன் பங்கீடு
(See stratified sampling)
Non-probability Sampling — வாய்ப்புத் தொடர்பற்ற மாதிரி முறை
Non response — பதிலின்மை
Non-sampling — மாதிரிமுறை சாராத
error — பிழை
variance — மாறுபாடு
Normal curve — இயல்நிலை வளைவு
distribution — பரவல்
population — முழுத்தொகுதி
Not-at-home — காண இயலாத
Numerator — பகுதி

O

Object — பொருள், நோக்கம்
Observation — கண்டறிதல்
Observed data — கண்டறிந்த விவரங்கள்
Optimum allocation — உத்தமப் பங்கீடு
Order — படி, வரிசை
Origin — ஆதி
Output — ஆக்கம், வெளிப்பாடு

P

Pair — இணை
Parameter — முழுத் தொகுதி அளவை
Part — பகுதி
Partial correlation — பகுதி ஒட்டுறவு
Percentage — சதவீதம்
Period — காலவட்டம்
Periodicity — காலவட்ட ஒழுங்கு

Periodic variation	—	திரும்பவரும் மாறுபாடு
Phase	—	நிலை, முகம்
Pilot enquiry	—	முன்னோடி கணிப்பு
survey	—	ஆய்வு
Population	—	முழுமைத் தொகுதி
census	—	மக்கட் தொகைக் கணிப்பு
finite	—	வரம்புடைத் தொகுதி
hypothetical	—	கற்பனைத் தொகுதி
infinite	—	வரம்பற்றத் தொகுதி
real	—	உண்மைத் தொகுதி
Plan	—	திட்டம்
sampling	—	மாதிரித் திட்டவரை
Precise	—	திட்ப
Precision	—	திட்பம்
Primary	—	முதல் நிலை
data	—	விவரங்கள்
Principle	—	தத்துவம்
Probability	—	நிகழ்தகவு, வாய்ப்பு
density	—	அடர்த்தி
function	—	சார்பு
normal	—	இயல் நிலை வாய்ப்பு
sampling	—	நிகழ்தகவு மாதிரிமுறை
proportional to size	—	அளவோடு வீதத் தொடர்புடைய வாய்ப்பு
Problem	—	பிரச்சினை, கணக்கு
Probable error	—	ஊகப் பிழை
Process	—	செயல்முறை, நடைமுறை
Proof	—	நிரூபணம்
Property	—	பண்பு
Proportion	—	வீதம், விகிதசமம்
estimate of	—	வீத மதிப்பீடு
Proportional allocation	—	வீதப் பங்கீடு (விகித அளவில் பங்கீடு)
Public opinion survey	—	மக்கள் கருத்து ஆய்வு
Punch card	—	துளை அட்டை
Purposive sampling	—	நோக்கமுள்ள மாதிரித் தேர்வு

Q

Qualitative	—	பண்பின்
Qualitative data	—	பண்பின விவரம்
Quality control	—	தரக் கட்டுப்பாடு
Quantity	—	கணியம், அளவு

Questionnaire — வினுப் பட்டியல்
Quota sampling — ஒதுக்கீட்டு மாதிரித் தேர்வு
Quinquennial — 5 ஆண்டுகட்கு ஒருமுறை

R

Random — ராண்டம்
sample — சமவாய்ப்பு மாதிரி
numbers — ராண்டம் எண்கள்
fluctuations — இயைபிலா மாறுபாடுகள்
variable — ராண்டம் மாறி (வாய்ப்பு) மாறி
Rate — வீதம்
birth — பிறப்பு வீதம்
death — இறப்பு வீதம்
employment — வேலைவாய்ப்பு வீதம்
Ratio — விகிதம்
estimation — விகித மதிப்பீடு
(See Ratio estimate)
Raw data — சீர்படா விவரம்
Real — உண்மையான
population — முழுத்தொகுதி
Record — பதிவேடு, குறிப்பேடு
Registration — பதிவு
sample — மாதிரிப் பதிவீட்டு முறை
Reasoning — ஆய்வு
Reference — குறிப்பீடு
period — குறிப்பீட்டுக் காலம்
Regression — மாறிகளின் தொடர்புப் போக்கு
coefficient — கெழு
analysis — பகுப்பாய்வு
estimate — மதிப்பீடு
linear — ,நேர்கோட்டு
Relative — சார்ந்த, ஒப்பீட்டு
deviation — விலக்கம்
precision — திட்பம்
Reliability — நம்பகத்தன்மை
Representative — பிரதிநிதி
sample — மாதிரி
Repetitive — திரும்பச் செய்யும்
Reporting period — அறிவிக்கும் காலம்
Report — அறிக்கை

Response	—	பதில்
variance	—	மாறுபாடு
Respondent	—	பதில் தருவோர்
Root mean square deviation	—	வர்க்கமூல வர்க்கச் சராசரி விலக்கம்
Roughly	—	சுமாராக

S

Sample	—	மாதிரி, கூறு
,average	—	சராசரி
,check	—	சரிபார்த்தால்
,design	—	திட்ட அமைப்பு
,observation	—	விவரம்
,random	—	ராண்டம் மாதிரி
,size	—	மாதிரி அளவு
,survey	—	மாதிரி ஆய்வு
,unit	—	மாதிரி அலகு
Sampling	—	மாதிரிமுறை, கூறெடுத்தல்
,average	—	கூறுகளின் சராசரி
,distribution	—	கூற்று பரவல் (மாதிரி)
,double	—	இருபடி மாதிரிமுறை
,error	—	மாதிரிமுறை சார்ந்த பிழை
,fluctuation	—	மாதிரிமுறை ஏற்றத்தாழ்வுகள்
,methods	—	மாதிரிமுறை வழிமுறைகள்
,multiple	—	பலபடி மாதிரிமுறை
,multiphase	—	பல்தோற்ற மாதிரிமுறை
,multistage	—	பலகட்ட மாதிரிமுறை
,probability	—	வாய்ப்பு மாதிரிமுறை
,purposive	—	நோக்கமுள்ள மாதிரிமுறை
,quota	—	ஒதுக்கீட்டு மாதிரி முறை
,random	—	ராண்டம் மாதிரிமுறை
,simple random	—	எளியராண்டம் மாதிரிமுறை
,units	—	அலகுகள்
,fraction	—	பின்னம்
stratified	—	படுகை மாதிரிமுறை
,systematic	—	ஒழுங்கு மாதிரிமுறை
,cluster	—	திரள் மாதிரிமுறை
,PPS	—	PPS மாதிரிமுறை
,judgement	—	தீர்ப்பு மாதிரிமுறை
Schedule	—	வினாத் தொகுதி (பட்டியல்)
Secondary data	—	வழிநிலை விவரங்கள்

Series	—	தொடர்வரிசை
Significance	—	முக்கியத்துவம்
Source	—	மூலம், தோற்றுவாய்
Solution	—	தீர்வு
Stable	—	உறுதியான, நிலையான
Standard	—	தரம்
deviation	—	விலக்கம்
error	—	பிழை
unit	—	அலகு
Statistic	—	மாதிரி அளவை
Statistical	—	புள்ளியியல் சார்ந்த
inference	—	ஊகம்
bureau	—	புள்ளி விவரச் செயலகம்
distribution	—	புள்ளியியல் பரவல்கள்
measures	—	அளவைகள்
methods	—	வழிமுறைகள்
quality control	—	தரக் கட்டுபாடு
reasoning	—	கருத்துமுறை
regularity	—	புள்ளி விவர ஒழுங்கு நியதி
research	—	புள்ளியியல் ஆராய்ச்சி
Statistics	—	புள்ளியியல், புள்ளிவிவரம்
official	—	அரசாங்கப் புள்ளிவிவரம்
agricultural	—	விவசாயப் புள்ளிவிவரம்
industrial	—	தொழில் புள்ளிவிவரம்
population	—	மக்கட்தொகைப் புள்ளிவிவரம்
trade	—	வர்த்தகப் புள்ளிவிவரம்
vital	—	பிறப்பிறப்புப் புள்ளிவிவரம்
Indian	—	இந்தியப் புள்ளிவிவரம்
Stratum	—	படுகை
Stratified	—	படுகை அமைத்த
Strata	—	படுகைகள்
Stratified sampling	—	படுகை அமைத்த மாதிரிமுறை
Stratification	—	படுகை பிரித்தல்
Sub-population	—	துணைத் தொகுதி
Super population	—	பெரும் முழுமைத் தொகுதி
Supplement	—	பிற்சேர்க்கை
Survey	—	அளவெடுப்பு
System	—	முறை
Systematic	—	முறையான
sampling	—	ஒழுங்கு மாதிரிமுறை

T

Table	—	பட்டியல், அட்டவணை
of random numbers	—	ராண்டம் எண் பட்டியல்
,contingency	—	நிகழ்வு அட்டவணை
Tabulate	—	பட்டியல் அமை
Tabulation	—	பட்டியலமைத்தல்
error	—	பிழை
,mechanical	—	பொறிவழி பட்டியலமைத்தல்
Tally mark	—	சரிபார்க்கும்குறி, கணிப்புக்குறி
Techniques	—	வழிமுறைகள்
,sampling	—	மாதிரித் தேர்வு வழிமுறைகள்
Tendency	—	போக்கு, இயல்பு
Test of significance	—	சிறப்புகாண் சோதனை
Time series	—	காலத்தொடர் வரிசை
Theoretical	—	அறிமுறைக்குரிய
Total	—	மொத்தம்
,population	—	முழுமைத் தொகுதி மொத்தம்
Trait	—	பண்பு
Trend	—	போக்கு
seasonal	—	காலப் போக்கு
cyclical	—	சுழற்சிப் போக்கு
Trade Statistics	—	வர்த்தகப் புள்ளிவிவரங்கள்
Trial	—	சோதனை

U

Ultimate	—	இறுதிநிலை
Ultimate cluster	—	இறுதிநிலைத் திரள்
Unbiased	—	பிறழ்ச்சியற்ற
Unbiased estimate	—	பிறழ்ச்சியற்ற மதிப்பீடு
Uniform	—	ஒரு சீரான
Unit	—	அலகு, கூறு
sampling	—	மாதிரிக் கூறு
Universe	—	பேரண்டம்
Universal	—	பொதுப்படையான
Unlimited	—	எல்லையற்ற
Unstable	—	நிலையற்ற
Upper Confidence Limit	—	மேல் நம்பிக்கை எல்லை
Utility	—	பயன்பாடு

V

Valid	—	ஏற்கக்கூடிய
Value	—	மதிப்பு
maximum	—	பெருமமதிப்பு

Value minimum — சிறுமதிப்பு
Variable — மாறி
,auxiliary — துணை மாறி
,dependent — சார்ந்த மாறி
,independent — தனித்த மாறி
,continuous — தொடர் மாறி
,random — ராண்டம் (வாய்ப்பு) மாறி
Variate — மாறி
,bi — இரு மாறி
,uni — ஒற்றை மாறி
,multi — பல் மாறி
Variation — மாறுபாடு
seasonal — பருவகால மாறுபாடு
cyclical — சுழல்மாறுபாடு
,coefficient of — மாறுபாட்டுக்கெழு
,between — மாறுபாட்டுக் இடையீட்டு மாறுபாடு

W

Weight — எடை, நிறை
Weighted mean — நிறையிட்ட சராசரி
index — குறியீடு
Wholesale — மொத்தவிற்பனை
index — குறியீடு
Work — பணி, வேலை
Worker — பணியாள்
Width — இடைவெளி
Within Variation — உள்ளிட்ட மாறுபாடு
With replacement — ஈடுசெய்யப்பட்ட
Without replacement — ஈடுசெய்யப்படாத

X

X-axis — X-அச்சு

Y

'Yes' or 'No' type — 'ஆம்' அல்லது 'இல்லை' வகை

Z

Zero — பூச்சியம்
Zones — வட்டங்கள் (பகுதிகள்)
,customs — சுங்கப் பகுதிகள்